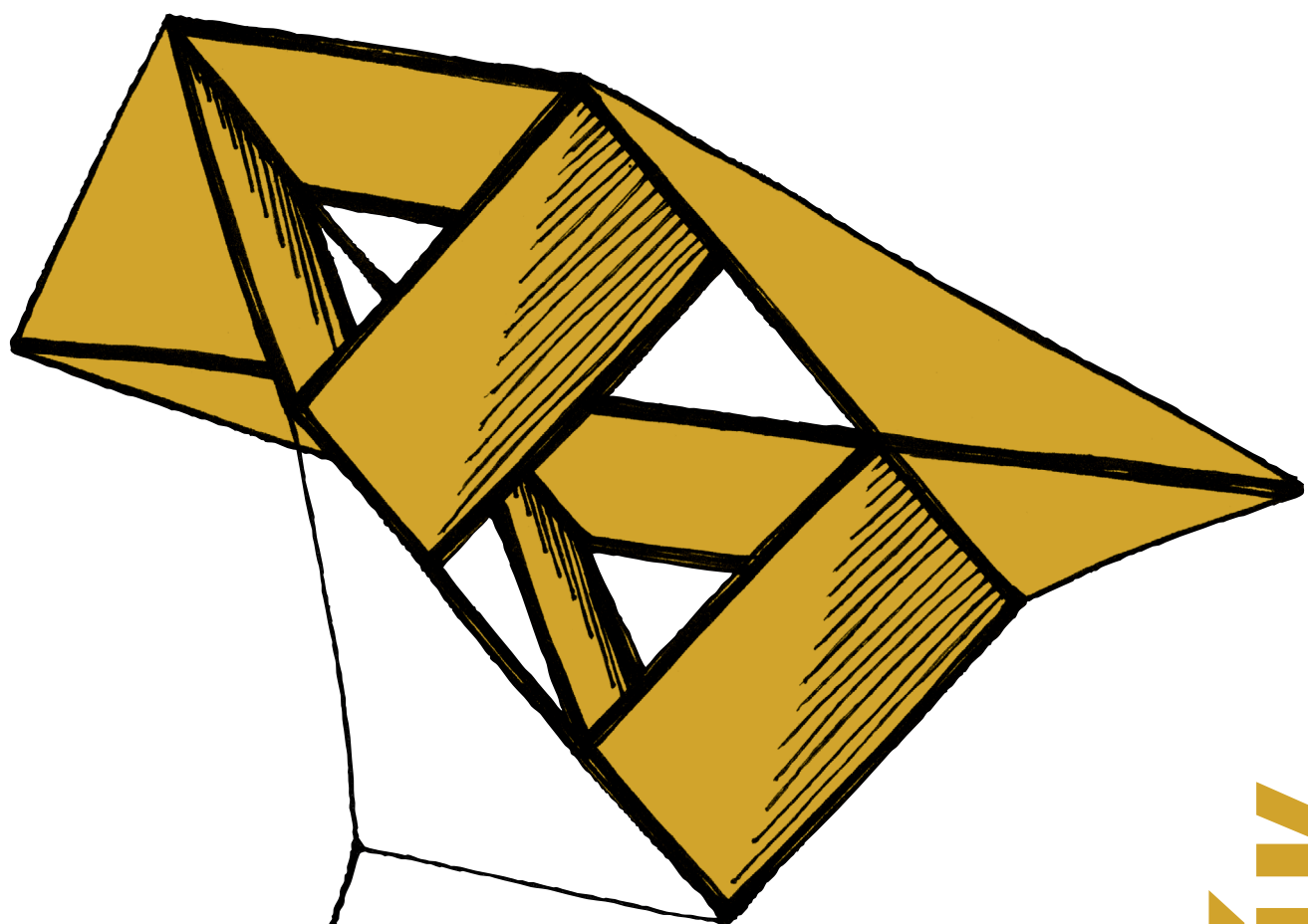
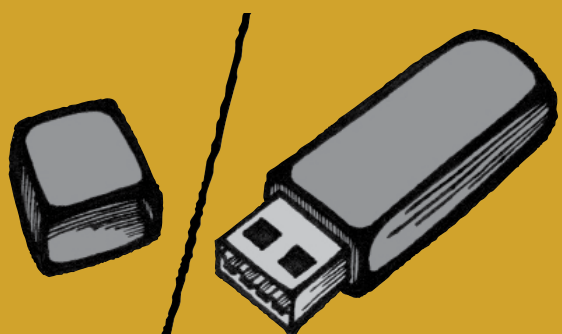




FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

SBORNÍK

STUDENTSKÁ VĚDECKÁ KONFERENCE



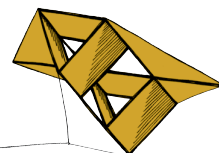
**MAGISTERSKÉ A DOKTORSKÉ
STUDIJNÍ PROGRAMY**

22.5.2025



FAKULTA APLIKOVANÝCH VĚD
ZÁPADOČESKÉ UNIVERZITY
V PLZNI

STUDENTSKÁ
VĚDECKÁ KONFERENCE **2025**



Magisterské a doktorské studijní programy

Sborník rozšířených abstraktů

Název: SVK FAV 2025 – magisterské a doktorské studijní programy

Editor: Jan Rendl

Ilustrace na obálce: Jana Trávníčková

Vydavatel: Západočeská univerzita v Plzni, Univerzitní 8, 306 14 Plzeň

Datum vydání: květen 2025

ISBN 978-80-261-1302-7

Tato publikace neprošla redakční ani jazykovou úpravou.

Sponzoři konference



Obsah

Sekce – Fyzika, matematika, geomatika, mechanika

Reduced-order modelling of blood flow in the human arterial tree Augustiňák Tomáš	8
Modelování kmitání rotorů s pasivními hltiči kmitů Bareš František	10
Global gravitational field modelling for spheroidal planetary bodies: non-singular expressions Belinger Jiří	12
Globální modelování gravitačních polí sféroidálních planetárních těles: studium Legendreových funkcí Dohnalová Veronika	14
Strongly thermochromic W-doped VO₂ films with large temperature coefficient of electrical resistance near room temperature Farrukh Sadoon	16
Ultra-low-resistivity nitrogen-doped Cu₂O thin films fabricated by reactive high-power impulse magnetron sputtering (HiPIMS) Koloros Jan	18
Deposition of Cu-Doped TiO₂ using Helium-Assisted DC Magnetron Sputtering for Conductometric Hydrogen Sensing Kumar Akash	20
Advancing Cu₂O Films for PEC Water Splitting Using Reactive HiPIMS: Enhanced Crystallinity and Photocurrent Performance Vosejpka Jan	22
Synthesis and characterization of HiPIMS-deposited Zr–O–N films for potential water splitting applications Vu Minh Thanh	24

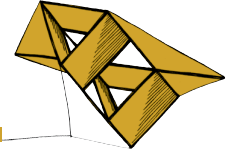
An Analytical Design Method for Power System Stabilizers Based on Hinf Specifications	
Brabec Michal	27
Monitorovací systém pro včelaře	
Březina Pavel	29
Zpracování obrazu pro automatické stříhání větví	
Gaier Jakub	31
Detekce zrcadlové symetrie v 3D geometrických modelech proměnných v čase	
Gregor Vít	33
Measurement Dimensionality and Its Impact on Position Error Covariance in Tracking Applications	
Gulyj Michael	35
Do Compact Multilingual Large Language Models Speak Czech?	
Hejman Jakub	37
Autoregressive Upscaling of Sparse Single–Cell Data Improves Interpretability	
Honzík Tomáš	39
Modul pro statistickou analýzu medicínských dat v systému Czech Salivary Gland Database	
Jelínek Vojtěch	41
Burzovní simulátor pro trh řízený limitními objednávkami	
Kimlová Vladimíra	43
Software for Neurorehabilitation	
Kment Tomáš	45
Detekce a prevence driftu konfigurace síťových komponent v síti Webnet.	
Koldovská Aneta	47
The Right Tracker for the Right Job: Maritime Case Study	
Krejčí Jan	49
Engineering Thermostable Proteins Using Large Language Models	
Kuhajda Lukáš	51
Analysis of White Matter Diffusion Properties in the Context of Selected Brain Tumors	
Kukrál Martin	53
Zero Vibration Filter Design Tool	
Langmajer Martin	55

Spacial Analysis of Image Description for Detection of Cognitive Disorders Lebeda Tomáš	57
Improving Sign Language Translation through Multimodal Language Alignment Majer Filip	59
Two Pillars of Software Project Health: DevOps and DevEx Majidov Elgun	61
Testing Platform for Periodic Control Myslivec Tomáš	63
SpiTranNet: A Spiking–Transformer network for Motor Imagery classification using multi–channel EEG signals Pham Duc Thien	65
Segmentace cév prasete domácího ve snímcích z výpočetní tomografie Romová Jana	67
Computer Vision Applications in Video Recordings for Traffic Signal Detection and Classification on Czech Railways Schnurpfeil Daniel	69
Modelování a identifikace portálových jeřábů Sukovatý Daniel	71
Klasifikace akustických událostí pro diagnostiku výrobní linky Šimek Jan	73
Cross–lingual Emotion Detection Šmíd Jakub	75
Využití transformerů pro specifickou úlohu z praxe Tauš Daniel	77
Point–mass Filter with Non–equidistant Grid Design Třebal Jan	79
Nonsense Word Repetition for Detection of Cognitive Disorders Tupý Jan	81
Vývoj autonomní mobilní platformy pro přepravu nákladů Tvrz Jakub	83
Using Pre–trained Models for Phoneme Representation in Czech Speech Synthesis Vladař Lukáš	85
Towards Aesthetic Enrichment of Mirror Selfies via Automatic Image Analysis Vyskočil Jiří	87

Complexity of KIR allele identification pipeline design Wolf Kateřina	89
Detekce anomálií v datech z knih limitních objednávek Zappe Dominik	91
Ablations Studies in Pose-Based Sign Language Translation Železný Tomáš	93

Sekce

Fyzika, geomatika, matematika, mechanika



Reduced-order modelling of blood flow in the human arterial tree

Tomáš Augustiňák¹

1 Introduction

Haemodynamics is recognised to affect the pathogenesis of the most life-threatening cardiovascular diseases, such as heart attack or brain stroke. Thus, in the new millennium, the need to better understand the complexity and uniqueness associated with pulsatile blood flow in compliant arteries has given rise to mathematical models of different spatial dimensions, making it possible to obtain accurate patient-specific results at low computational cost.

2 Description of the model

This work presents a well-known model of the arterial system consisting of 55 largest arteries in the human body, presented in Sherwin et al. (2003), Fig. 1. In this work, it is solved by the MacCormack finite volume method, whereas in the literature the authors used the discontinuous Galerkin finite element method. A comparison of computed results enables their verification.

The model comprises of individual arterial segments formulated as one-dimensional models. This formulation employs certain simplifying assumptions: The vascular wall is assumed thin-walled, radially symmetric, and subject to small deformations, so one can use Hooke's law of linear elasticity. Blood is assumed homogeneous and incompressible, and pressure p and velocity U are constant in a given lumen cross-section A . Wall inertia is neglected. The model can be formulated as a hyperbolic system of partial differential equations

$$\frac{\partial}{\partial t} \begin{bmatrix} A \\ U \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} AU \\ \frac{U^2}{2} + \frac{\beta}{\rho} (\sqrt{A} - \sqrt{A_0}) + \frac{p_{\text{ext}}}{\rho} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (1)$$

where x is the spatial coordinate, t is time, ρ is blood density, and β relates to wall stiffness.

The individual segments are connected into a branching structure using a set of six conditions in each bifurcation, the solution of which gives a correction to the values on the boundaries. Three of these conditions express mass conservation in the bifurcation and total pressure continuity in all three vessels. The other three conditions are imposed on the so-called characteristic variables. The perturbation of these variables can only move into the bifurcation and are assumed to be unaffected by the computed correction.

Finally, one prescribes boundary conditions. At the inlet, a boundary condition is given as a value of the forward characteristic variable, while the backward one is approximated from the previous time level. This allows for the calculation of primary unknowns, A and U . At the outlets, a terminal resistance is prescribed, that gives a ratio of characteristic variables. The forward characteristic is approximated from known data, the other one can be easily calculated from the terminal resistance, so primary unknowns can be computed.

¹ student of the doctoral degree program Applied Mechanics, field of study Mechanics, e-mail: tau-gusti@ntis.zcu.cz

3 Results and concluding notes

Firstly, from the comparison with Sherwin et al. (2003), one can clearly see a very good correspondence between the computed and reference results. Secondly, even though this model does not contain any representation of the heart and even though outlet boundary conditions are very simple, one can spot the dicrotic notch, that is the local minimum of lumen cross-section in the ascending aorta, in the results for this vessel. Finally, the model can be regarded as multiscale because the terminal resistance corresponds to the one-element Windkessel model (a resistor) connected to the outlet. Terminal resistances significantly impact results in the entire arterial tree and, as such, will play a pivotal role in the personalisation of the model.

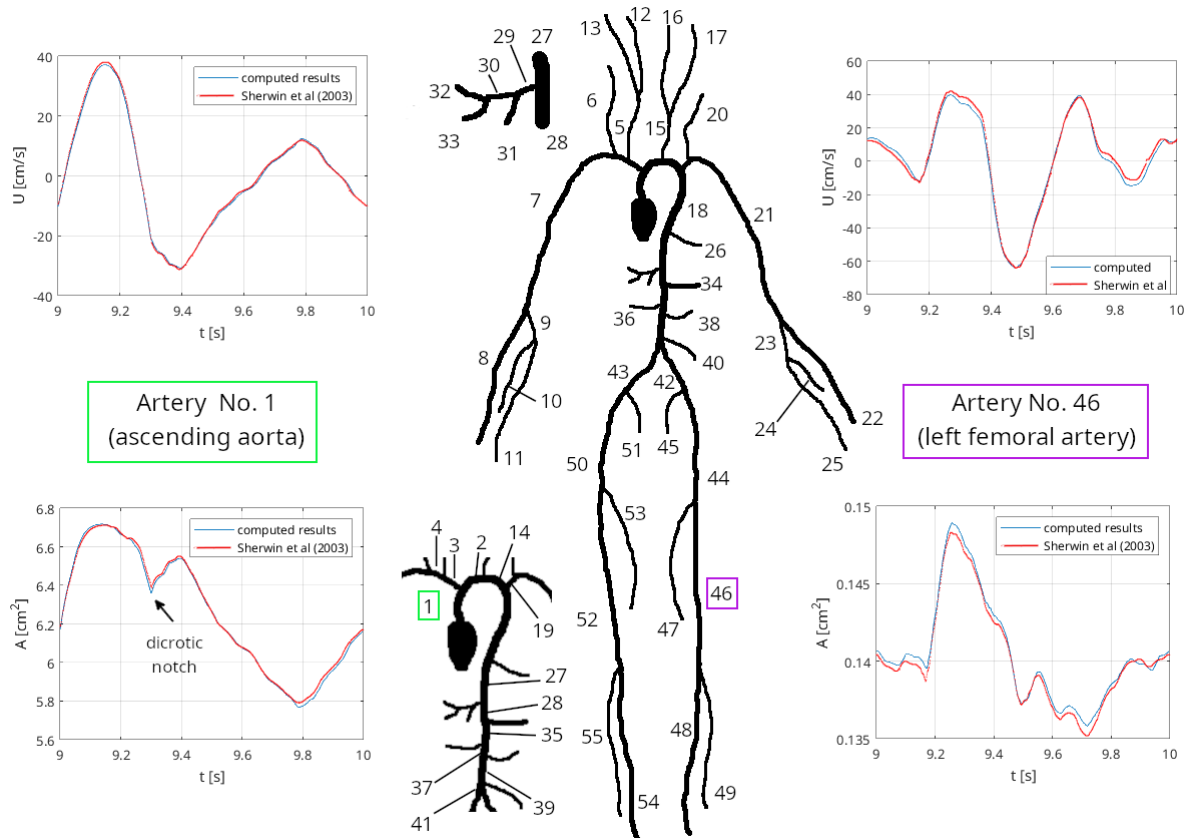


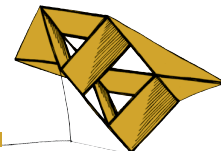
Figure 1: Illustration of 55 arteries modelled in this work

Acknowledgement

The work was supported by the internal grant project SGS-2025-015 of the University of West Bohemia.

References

Sherwin, S. J., Formaggia, L., Peiró, J. and Franke, V. (2003) Computational modelling of 1D blood flow with variable mechanical properties and its application to the simulation of wave propagation in the human arterial system. *International Journal for Numerical Methods in Fluids*, Vol. 43, pp. 673–700.



Modelování kmitání rotorů s pasivními hltiči kmitů

František Bareš¹

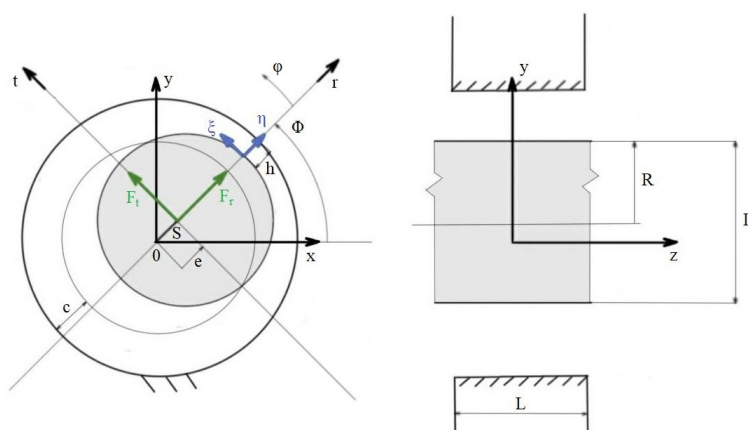
1 Úvod

Tato práce se zabývá tvorbou diskrétních modelů rotorové soustavy a navržením parametrů pasivního hltiče kmitů pro účinné snížení vibrační hřídelového čepu v rezonanční oblasti provozu. Zejména je kladen důraz na ovlivnění chování rotoru hydrodynamickými silami způsobené mazivem v kluzných ložiscích. Tyto síly mají negativní vliv na chod systému zejména při malé hmotnosti rotorové části.

2 Hydrodynamické síly

Hydrodynamické síly byly popsány na základě Reynoldsovy rovnice. Aby bylo možné explicitně popsat tyto síly jako funkci polohy a rychlosti čepu, byl využit model takzvaně nekonečně krátkého ložiska. Tento předpoklad se opírá o to, že poměr délky a vnitřního průměru ložiska nepřekračuje stanovenou hodnotu. Složky sil v polárním souřadnicovém systému jsou určeny vztahy [1]

$$\begin{aligned} F_r^{IS}(\varepsilon, \dot{\varepsilon}, \dot{\Phi}, \omega) &= -\mu R L \left(\frac{L}{c}\right)^2 \left[|\omega - 2\dot{\Phi}| \frac{\varepsilon^2}{(1 - \varepsilon^2)^2} + \frac{\pi \dot{\varepsilon} (1 + 2\varepsilon^2)}{2(1 - \varepsilon^2)^{5/2}} \right], \\ F_t^{IS}(\varepsilon, \dot{\varepsilon}, \dot{\Phi}, \omega) &= \mu R L \left(\frac{L}{c}\right)^2 \left[(\omega - 2\dot{\Phi}) \frac{\pi \varepsilon}{4(1 - \varepsilon^2)^{3/2}} + \frac{2\varepsilon \dot{\varepsilon}}{(1 - \varepsilon^2)^2} \right]. \end{aligned} \quad (1)$$



Obrázek 1: Zavedené parametry kluzného ložiska.

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

3 Určení parametrů hltiče

Pro optimalizaci parametrů hltiče bylo nutné vyjádřit hydrodynamické síly (1) vyjádřit ve stacionárním případě, tedy byla zanedbána rychlost pohybu hřídelového čepu [2]. Díky tomuto vyjádření a převedení složek sil do kartézského souřadnicového systému bylo možné určit matice tuhosti $\mathbf{K}^{HD}$ a tlumení $\mathbf{B}^{HD}$. Ty byly dosazeny do modelu rotou daného vztahem

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{K}\mathbf{q} + \mathbf{B}\dot{\mathbf{q}} = \mathbf{f}_e + \mathbf{f}_g. \quad (2)$$

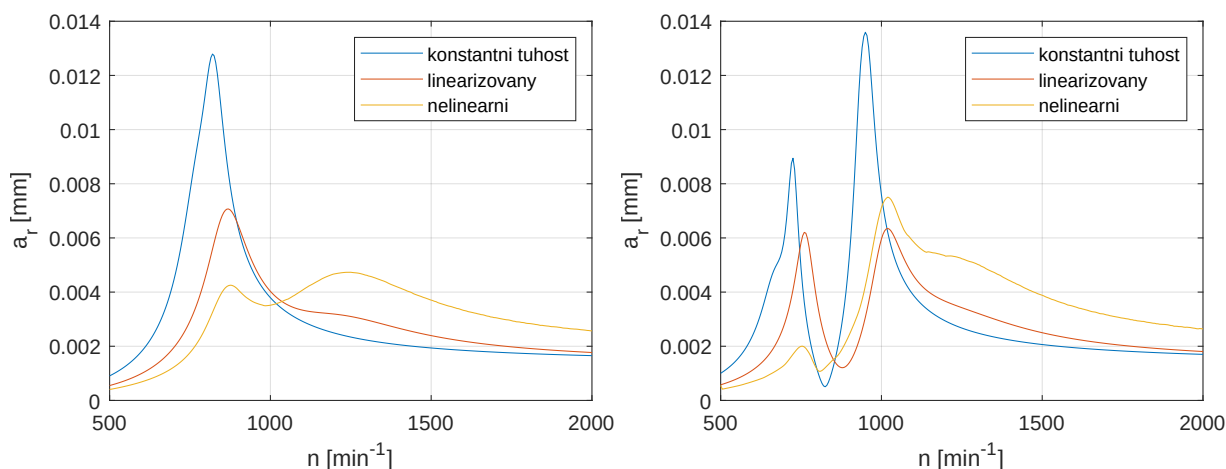
Jako parametr pro optimalizaci parametrů hltiče byla použita amplituda hřídelového čepu. Amplitudy systém byly určeny ze soustavy rovnic [3]

$$\underbrace{(\mathbf{K} - \omega^2 \mathbf{M} + i \omega \mathbf{B})}_{\mathbf{Z}(\omega)} \tilde{\mathbf{q}} = \tilde{\mathbf{f}}, \quad (3)$$

kde $\mathbf{Z}$ je matice dynamické tuhosti.

4 Výsledky

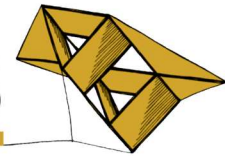
Byly porovnány tři výsledky, které se lišily v přístupu vyjádření hydrodynamických sil v ložiscích. Nejprve byly koeficienty tuhosti a tlumení konstantní. Následně byly uvažovány závislé na úhlové rychlosti rotoru. Koeficienty byly tedy dány stacionárním případem vztahu (1). Nakonec byly získané parametry hltiče dosazeny do modelu s nestacionárním případem hydrodynamických sil. Ve všech třech případech byla redukována amplituda hřídelového čepu v rezonanční oblasti.



Obrázek 2: Porovnání vypočtených amplitudových charakteristik bez aplikovaného hltiče a s ním.

Literatura

- [1] Y. Bastani and M. de Queiroz. A new analytic approximation for the hydrodynamic forces in finitlength journal bearings. *Journal of Tribology*, 2010.
- [2] Y. Hori. *Hydrodynamic Lubrication*. Springer-Verlag Tokyo, 2002.
- [3] Slavík, Stejskal, and Zeman. *Základy dynamiky strojů*. Vydavatelství ČVUT, Praha, 1997.



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@ntis.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).

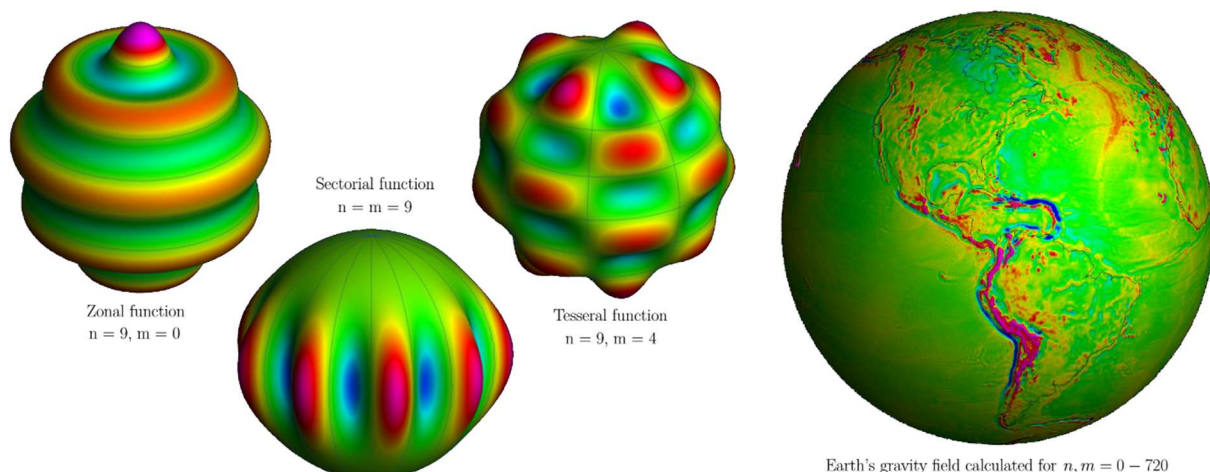


Figure 1: Synthesis of Earth's gravity field from a large set of spherical harmonic functions (spatial resolution rises with degree n and order m) – Ince et al. (2019).

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

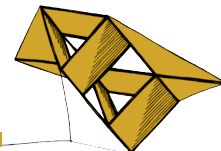
This contribution describes aspects of the gravitational field modelling of planetary bodies using spheroidal harmonic synthesis. We focus on the substitution of singular expressions (containing partial derivatives with respect to reduced spheroidal latitude β and trigonometric functions dependent on β) occurring in calculations by non-singular recursive expressions. Similar approach was already tested for spherical approximation by Hamáčková et al. (2016)

Acknowledgement

Authors were supported by the project No. 23-07031S of the Czech Science Foundation.

References

- Hamáčková E., Šprlák M., Pitoňák M., Novák P. (2016) Non-Singular Expressions for the Spherical Harmonic Synthesis of Gravitational Curvatures in a Local North-Oriented Reference Frame. *Computers & Geosciences* 88: 152-162. <https://doi.org/10.1016/j.cageo.2015.12.011>
- 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>



Globální modelování gravitačních polí sféroidálních planetárních těles: studium Legendreových funkcí

Veronika Dohnalová¹, Jiří Belinger², Martin Pitoňák³, Martin Pitoňák⁴, Pavel Novák⁵,
Michal Šprlák⁶

1 Úvod

Určování gravitačních polí planetárních těles je významným tématem moderní geodézie. Zároveň je znalost gravitačních polí jádrem mnoha důležitých aplikací, jako je studium vnitřních struktur planetárních těles, stanovení planetárních referenčních systémů pro určování polohy, předpovídání oběžných drah umělých družic nebo navigace pozemních a kosmických vozidel.

Standardní koncepční rámec pro určování gravitačního pole ve fyzikální geodézii běžně vychází ze sférické aproximace. Nicméně současné výzkumy planetárních těles sluneční soustavy ukázaly, že mnohá z nich se podobají protáhlým nebo zploštělým elipsoidům Šprlák et al. (2020). Je tedy nutné formulovat moderní koncepční rámec, který využije aproximaci elipsoidickou.

2 Projekt elipsoidického modelování

V rámci projektu Elipsoidického modelování planetárních gravitačních polí byla formulována nová matematická teorie. Základní rovnicí, ze které projekt vychází, je rozvoj gravitačního potenciálu do řady sféroidálních harmonických funkcí uvedený v rovnici (1). Na základě tohoto vztahu byly odvozeny složky gravitačního gradientu, složky gravitačního tenzoru druhého řádu a složky gravitačního tenzoru třetího řádu.

$$V(u, \Omega) = \frac{GM}{R_o} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \frac{Q_{n,|m|} \left(i \frac{u}{\varepsilon} \right)}{Q_{n,|m|} \left(i \frac{b}{\varepsilon} \right)} \bar{C}_{n,m}^o(u) \bar{Y}_{n,m}(\Omega) + \frac{GM}{R_i} \sum_{n=0}^{\infty} \sum_{m=-n}^{+n} \frac{P_{n,|m|} \left(i \frac{u}{\varepsilon} \right)}{P_{n,|m|} \left(i \frac{b}{\varepsilon} \right)} \bar{C}_{n,m}^i(u) \bar{Y}_{n,m}(\Omega). \quad (1)$$

Software pro výpočet gravitačního potenciálu byl zatím publikován pro vnější rozvoj.

¹ student navazujícího studijního programu Geomatika, specializace Geoinformatika, e-mail: dohnalv@ntis.zcu.cz

² NTIS – Nové technologie pro informační společnost, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, e-mail: belinger@ntis.zcu.cz

³ NTIS – Nové technologie pro informační společnost, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, e-mail: pitonakm@kgm.zcu.cz

⁴ NTIS – Nové technologie pro informační společnost, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, e-mail: pitonakm@kgm.zcu.cz

⁵ NTIS – Nové technologie pro informační společnost, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, e-mail: panovak@ntis.zcu.cz

⁶ NTIS – Nové technologie pro informační společnost, Fakulta aplikovaných věd, Západočeská univerzita v Plzni, e-mail: sprlakm@kgm.zcu.cz

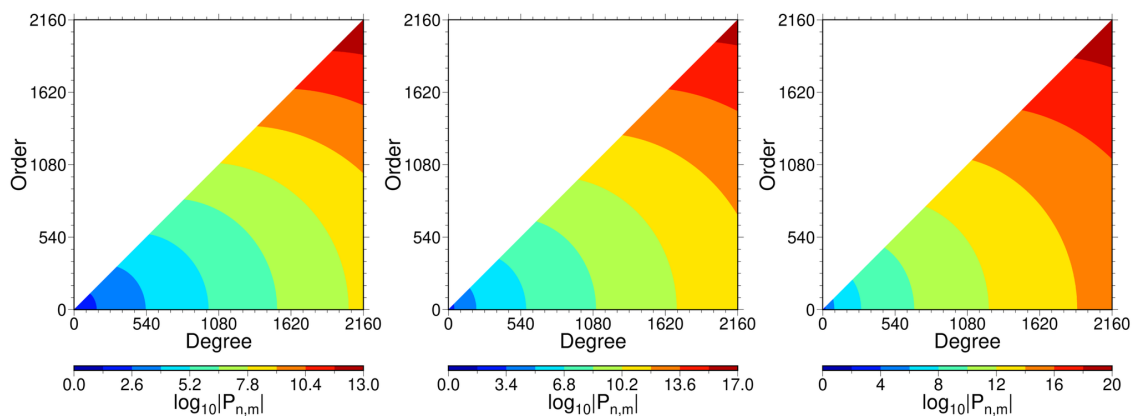
Nyní se projekt zaměřuje na doplnění softwaru o vnitřní rozvoj a o výpočet normovaných sféroidálních harmonických koeficientů $\bar{C}_{n,m}^o(u)$ a $\bar{C}_{n,m}^i(u)$.

3 Studium Legendreových funkcí

Významnou součástí rozvoje (1) jsou Legendreovy funkce prvního druhu $P_{n,|m|}$ a druhého druhu $Q_{n,|m|}$. Běžně se k jejich výpočtu používají rekurze, které umožňují rychlý výpočet, ale mohou numericky selhat. Alternativní variantou je použití hypergeometrických funkcí.

Již v softwaru s vnějším rozvojem byly k výpočtu Legendreových funkcí využity funkce hypergeometrické. Jejich výpočet byl realizován pomocí knihovny GNU Scientific Library (GSL) Šprlák et al. (2020). Aby nový kompletní software nebyl závislý na použití cizí knihovny, byly v rámci projektu zkoumané různé hypergeometrické funkce a jejich vlastnosti.

Příspěvek se zaměřuje na výsledky z pozorování chování Legendreových funkcí v závislosti na různých vstupních parametrech jako je typ elipsoidu, výška výpočtového bodu a velikost zploštění. Dále byla sledována přesnost a numerická stabilita výpočtu. Zároveň byly zkoumané i první, druhé a třetí derivace a integrály Legendreových funkcí. Ty se objevují při výpočtu gradientu gravitačního potenciálu, jeho složek tenzorů druhého a třetího řádu a při výpočtu sféroidálních harmonických koeficientů.



Obrázek 1: Ukázka chování první (vlevo), druhé (uprostřed) a třetí (vpravo) derivace funkce $P_{n,|m|}$

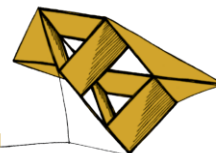
Korektnost výpočtu Legendreových funkcí prvního a druhého druhu byla prozatím otestována na zmíněném softwaru pro výpočet gravitačního potenciálu pomocí vnějšího rozvoje do řady sféroidálních harmonických funkcí.

Poděkování

Příspěvek byl podpořen grantovým projektem GAČR 23-07031S a projektem Západočeské Univerzity SGS-2022-027.

Literatura

Šprlák, M., Han, S.-C., Featherstone, W.E. (2020) Spheroidal forward modelling of the gravitational fields of 1 Ceres and the Moon. *Icarus*, Volume 335, Article 113412.



Strongly thermochromic W-doped VO₂ films with a large temperature coefficient of electrical resistance near room temperature

Sadoon Farrukh¹, Jaroslav Vlček², Jiří Rezek³, Radomír Čerstvý⁴, Tomáš Kozák⁵

1 Introduction

Vanadium dioxide (VO₂) is an extremely interesting and increasingly investigated coating material due to its reversible first-order transition between a low-temperature monoclinic VO₂(M1) semiconducting phase and a high-temperature tetragonal VO₂(R) metallic phase relatively near room temperature (approximately 68 °C for the bulk material). High modulation of the infrared transmittance, and electrical and thermal conductivity makes VO₂-based films a suitable candidate for numerous applications, such as electronic and optical switches, thermal sensors, smart thermal radiator devices for spacecraft, adaptive thermal camouflage, and energy-saving smart windows with automatically varied solar energy transmittance.

The application potential of these films depends on the ability to achieve not only the VO₂ stoichiometry but also the crystallization of the VO₂(M1/R) phase under as industry-friendly process conditions as possible, i.e., at a deposition temperature close to 300 °C (usually used temperatures are higher than 450 °C) and without any substrate bias voltage in case of usually used magnetron sputter techniques. Moreover, the transition temperature needs to be reduced to room temperature for many applications (e.g., energy-saving smart windows, new uncooled infrared detectors and high-performance temperature sensors). Besides the optical transmittance and the electrical resistivity below transition temperature, the characteristics of the semiconductor-metal transition, such as the corresponding phase-transition amplitude, hysteresis width and phase-transition sharpness, are of key importance.

Furthermore, VO₂ coatings can be combined with complementary functional layers to enhance durability and tailor the spectral response for specific environments, offering additional flexibility in device design without compromising core performance.

The paper deals with crystal structure, electronic band structure, optical and electrical properties, and semiconductor-metal transition characteristics of strongly thermochromic W-doped VO₂ films with a large (up to -12% K⁻¹) temperature coefficient of electrical resistance near room temperature (24 - 28 °C). They were deposited at a reduced temperature (350 °C) onto soda-lime glass substrates with two versions of Y-stabilized ZrO₂ (YSZ) interlayers (serving also as a highly optically transparent bottom antireflection layer) possessing different crystal orientations, and onto bare glass and monocrystalline YSZ and Al₂O₃ substrates for comparison.

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

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

³ Researcher at NTIS, field Plasma physics and physics of thin films, email: jrezek@kfy.zcu.cz

⁴ Researcher at NTIS, field Plasma physics and physics of thin films, email: cerstvy@kfy.zcu.cz

⁵ Associate Professor of physics, field Plasma physics and physics of thin films, NTIS, email: kozakt@kfy.zcu.cz

Additionally, the use of YSZ interlayers not only improves the optical impedance matching but also promotes better film adhesion and structural uniformity across the substrate surface.

The W-doped VO₂ films were synthesized using reactive deep oscillation magnetron sputtering (DOMS) with a pulsed O₂ flow control and to-substrate O₂ injection into a high-density plasma in front of the sputter V-W (3.0 wt.%) target. The DOMS is a modified version of high-power impulse magnetron sputtering with macropulses (500 μ s in this work) composed of short high-power micropulses (10 in this work) making it possible to increase discharge stability and deposition rate of films (20 – 30 nm min⁻¹ in this work).

2 Figures

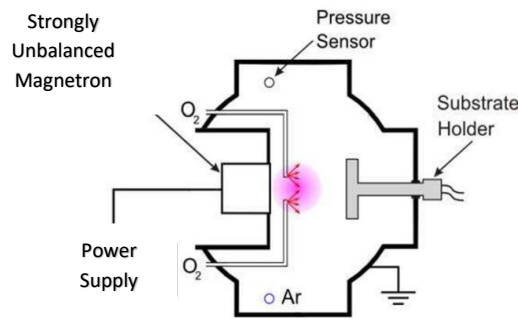
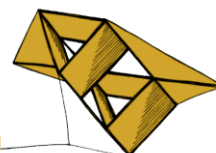


Figure 1: Schematic of deep oscillation magnetron sputtering (DOMS) with a pulsed O₂ flow control and to-substrate O₂ injection

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.



Ultra-low-resistivity nitrogen-doped Cu_2O thin films fabricated by reactive high-power impulse magnetron sputtering (HiPIMS)

Jan Koloros¹, Jiří Rezek², Pavel Baroch²

1 Introduction

Nowadays, transparent conductive oxides are widely used in many important applications. Whereas n-type transparent conducting oxides are well researched and widely used, there needs to be more **p-type** counterparts with sufficient performance. These **p-type TCOs** with comparable performance to their n-type counterparts are necessary to fabricate active electronic devices based on transparent **p–n junctions** (Singh et al., 2024). **Cu_2O -based** materials are a good candidate for this as they contain only abundant and non-toxic elements (Kumar et al., 2020). In this work, we systematically investigated the role of nitrogen incorporated in Cu_2O thin films on optical and electrical properties, namely optical band gap and electrical resistivity. We aim to prepare TCO layers with a high optical bandgap and low electrical resistivity with the highest possible mobility of free hole carriers.

The most suitable $\text{Cu}_2\text{O}:\text{N}$ films were prepared by reactive **HiPIMS** of Cu circular target (100 mm in diameter) in $\text{Ar}+\text{O}_2+\text{N}_2$ atmosphere. The pulse-averaged target power density (S_{da}) varied from ≈ 100 – 1300 Wcm^{-2} , and the fraction of N_2 in $(\text{Ar}+\text{N}_2)$ mass flow was 0–90 %. The optimal oxygen partial pressure was 260 mPa.

2 Result

The resistivity and optical bandgap of $\text{Cu}_2\text{O}:\text{N}$ films with the increasing amount of nitrogen measured using a four-probe method are shown in **Figure 1 (a)**, a decreasing trend can be seen with the increasing amount of nitrogen; the pulse-averaged target power density (S_{da}) is constant at 1300 Wcm^{-2} . **Figure 1 (b)** shows the effect of pulse-averaged target power density (S_{da}), the graph shows an increase of optical bandgap between 100 and 400 Wcm^{-2} , the increase may indicate a change in the crystal phase.

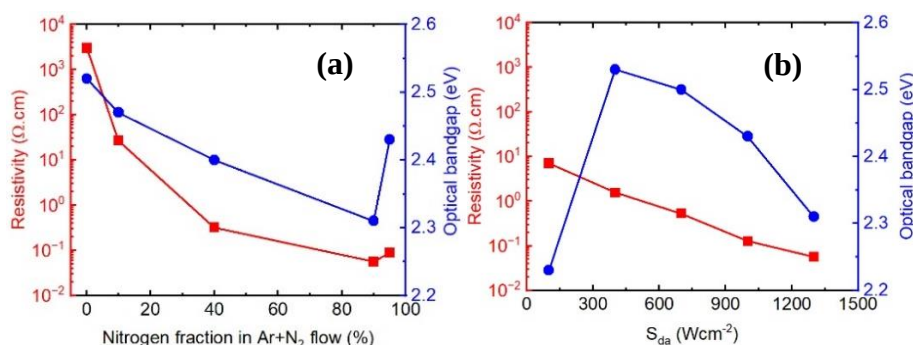


Figure 1: The resistivity and bandgap of $\text{Cu}_2\text{O}:\text{N}$ films (a) as a function of nitrogen fraction in $\text{Ar}+\text{N}_2$ mass flow (b) as a function of S_{da} with a constant nitrogen fraction of 90 %

¹ Ph.D. student of study program Plasma Physics and Physics of Thin Films, e-mail: koloros@ntis.zcu.cz

² Researcher, NTIS, VP4, Faculty of Applied Sciences, e-mail: jrezek@ntis.zcu.cz

In **Figure 2**, there are XRD patterns of two films, where for the films with the lowest $S_{da} = 100 \text{ Wcm}^{-2}$, the CuO phase is detected, and for the other layers with a different S_{da} , only the Cu₂O phase was detected. The film with the largest S_{da} was selected for comparison. This explains the significant increase in the optical bandgap $S_{da} = 100$ and 400 Wcm^{-2} in **Figure 1 (b)**. It also agrees with Balik (2019), which indicates that CuO has a lower optical bandgap than Cu₂O.

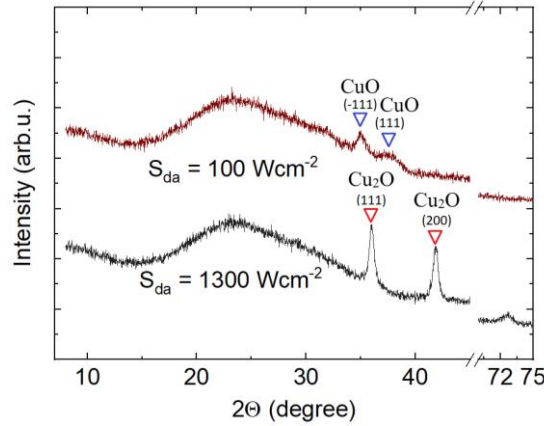


Figure 2: XRD of Cu₂O:N films with different pulse-averaged target power density S_{da}

3 Conclusions

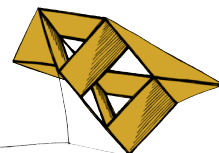
We have shown that reactive high-power impulse magnetron sputtering (**HiPIMS**) is a promising way for nitrogen-doped Cu₂O films. The film exhibits an electrical conductivity three orders of magnitude higher than a pure Cu₂O film for the optimized nitrogen concentration. Very **low resistivity** Cu₂O:N films with a resistivity of around $6 \times 10^{-2} \Omega\text{cm}$ were achieved. From XRD analysis and resistivity measurements, higher electrical conductivity and optical bandgap are achieved with the Cu₂O phase structure than CuO.

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

- J. Singh, P. Bhardwaj, R. Kumar, V. Verma, Progress in Developing Highly Efficient p-type TCOs for Transparent Electronics: A Comprehensive Review, J Electron Mater (2024). <https://doi.org/10.1007/s11664-024-11445-7>
- R. Kumar, K. Bergum, H. N. Riise, E. Monakhov, A. Galeckas, B. G. Svensson, Impact of post annealing and hydrogen implantation on functional properties of Cu₂O thin films for photovoltaic applications: Journal of Alloys and Compounds 825 (2020) 153982. <https://doi.org/10.1016/j.jallcom.2020.153982>
- M. Balik, V. Bulut, I. Y. Erdogan, Optical, structural and phase transition properties of Cu₂O, CuO and Cu₂O/CuO: Their photoelectrochemical sensor applications: International Journal of Hydrogen Energy (2019). <https://doi.org/10.1016/j.ijhydene.2018.08.159>



Deposition of Cu-Doped TiO_2 using Helium-Assisted DC Magnetron Sputtering for Conductometric Hydrogen Sensing

Akash Kumar¹, Nirmal Kumar², Stanislav Haviar³

1 Introduction

Hydrogen is a clean, abundant energy source crucial for zero-emission, hydrogen's high combustibility and wide explosive range in a mixture with oxygen (from 4 %) pose substantial safety risks [Patel et al. (1982), Dagdougui et al. (2018)], necessitating robust detection and monitoring systems. This study presents a novel helium-assisted sputter deposition method for Cu-doped TiO_2 conductometric sensors, intending to enhance the active surface area through microstructure modification, thereby improving sensing capabilities. This is achieved by careful optimization of magnetron sputtering, while the working gas consists of a mixture of argon, helium and oxygen. Partial replacement of argon with helium leads to the formation of structured porous thin films. The influence on the sensing response and the possible explanation of the effect in helium-rich working gas are discussed [Haviar et al. (2018), kumar et al. (2021)]

2 Methodology

Thin films were fabricated using conventional DC magnetron sputtering, utilizing a circular titanium target with an adjacent copper strip in controlled working gas mixtures of argon, helium, and oxygen, maintaining a total pressure of approximately 520 mPa. Helium was introduced as a replacement for part of argon in the working gas mixture of argon and oxygen and after that the glancing angle deposition (GLAD) technique was applied. Both Helium, GLAD were used to enhance the films' active surface area. The copper doping was chosen to make the films more conductive since the undoped titania films are nearly insulating. After deposition, samples were annealed at 400 °C in ambient air to crystallise and stabilize the films for hydrogen sensing. A series of different helium concentrations was prepared, varying the helium content from 0% to 95% in the working gas.

3 Result

X-ray diffraction patterns revealed the effect of helium on film microstructure. Films deposited under helium exhibit both anatase and rutile phases, whereas films deposited without helium show only the anatase phase. Scanning electron microscope (Figure 1(a)) observation reveals a grain size reduction with increasing plasma helium content. Nevertheless, the most pronounced effect was the distinct change in the structure, which appears highly porous for the films prepared with helium. Conductometric sensing was carried out using a four-point probe (Figure 1(b)), demonstrating that helium usage leads to a 6 times increase in sensor response for Cu-doped TiO_2 films (maximum He vs 0% Helium).

¹ student of the doctoral degree program Applied Sciences, Physics of Thin Films, e-mail: akash@ntis.zcu.cz

² Department of Physics NTIS, University of West Bohemia, e-mail: kumarn@ntis.zcu.cz

³ Department of Physics NTIS, University of West Bohemia, e-mail: haviar@ntis.zcu.cz

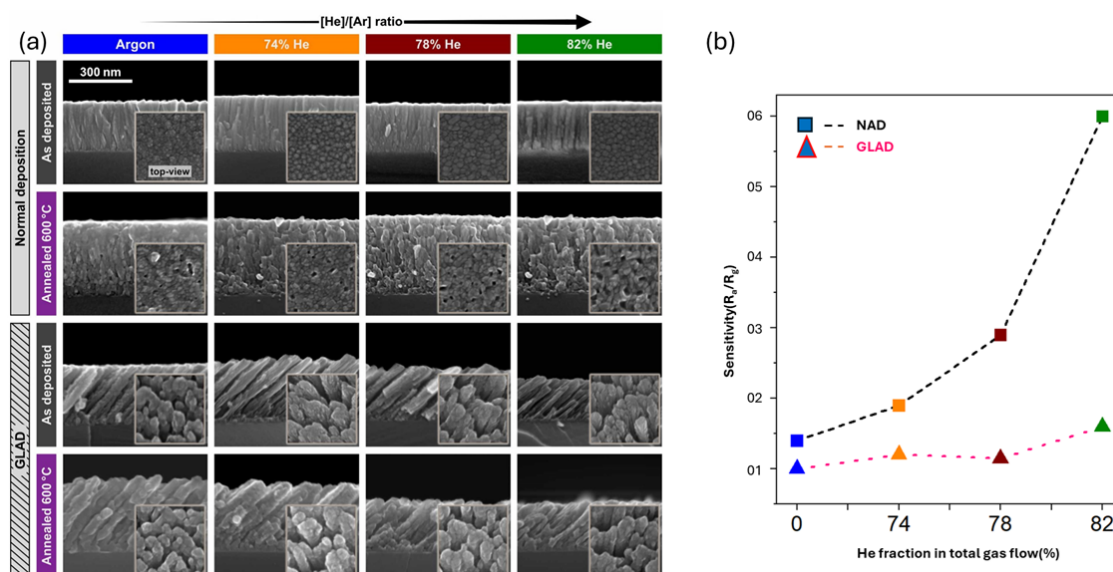


Figure 1: (a) SEM micrograph of NAD and GLAD Cu-doped TiO₂ in mixture of Ar, Helium(He) (b) sensitivity of prepared various He concentration sample

4 Conclusion

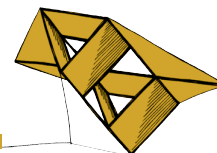
This study aimed to enhance hydrogen sensing of Cu-doped TiO₂, using two step modification strategies. The introduction of helium to the plasma significantly improved sensor sensitivity by increasing the active surface area. Thus, helium emerges as a highly effective method for optimizing hydrogen sensor performance.

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 Dr. Radomír Čerstvý, University of West Bohemia, for XRD measurements and analyses.

References

- Dagdougui, H., 2018. Hydrogen Logistics: Safety and Risks Issues. In: *Hydrogen Infrastructure for Energy Applications. 1st ed* Available at: <https://doi.org/10.1016/B978-0-12-812036-1.00007-X>
- Haviar, S., Capek, J., Batkov, S., Kumar, N., Dvořák, F., Duchon, T., Fialova, M. and Zeman, P., 2018. Hydrogen gas sensing properties of WO₃ sputter-deposited thin films enhanced by on-top deposited CuO nanoclusters. *International Journal of Hydrogen Energy*. Available at: <https://doi.org/10.1016/j.ijhydene.2018.10.127>
- S. V. Patel, J. L. Gland, J. W. Schwank, Film structure and conductometric hydrogen-gas-sensing characteristics of ultrathin platinum films, *Langmuir* 15 (1999) 3307–3311. doi:10.1021/la9809426.
- Kumar, N., Haviar, S. and Zeman, P., 2021. Three-Layer PdO/CuWO₄/CuO System for Hydrogen Gas Sensing with Reduced Humidity Interference. *Nanomaterials*, 11(3456), pp.1-16. Available at: <https://doi.org/10.3390/nano11123456>



Advancing Cu₂O Films for PEC Water Splitting Using Reactive HiPIMS: Enhanced Crystallinity and Photocurrent Performance

Jan Vosejpka¹, Jiří Čapek²

1 Introduction

Hydrogen is seen as an alternative, environmentally friendly energy carrier suitable for many purposes connected to decarbonisation. The main objective of current research is to develop a sufficiently cheap method of producing hydrogen so that it can compete economically with fossil fuels and replace them in the near future (IEA 2024). Photoelectrochemical (PEC) water splitting is an alternative to conventional water electrolysis. The process is based on a semiconductor that generates the redox potentials required for water splitting by photoexciting electron-hole pairs.

Cuprous oxide (Cu₂O) meets the theoretical requirements for efficient PEC water splitting, which makes it a key material for research in this field (Wick and Tilley 2015). However, conventional methods such as chemical synthesis and DC magnetron sputtering often fail to produce Cu₂O films with the desired density, crystallinity, and single-phase composition, i.e. properties essential for efficient water splitting.

High-Power Impulse Magnetron Sputtering (HiPIMS) is a cutting-edge technique for thin-film deposition. During each pulse, the discharge power density is typically two orders of magnitude higher than the average power density over the entire period. This results in a high level of ionization for both the working gas and the target material, along with an expanded ion energy distribution. These unique features enable enhanced control over film properties.

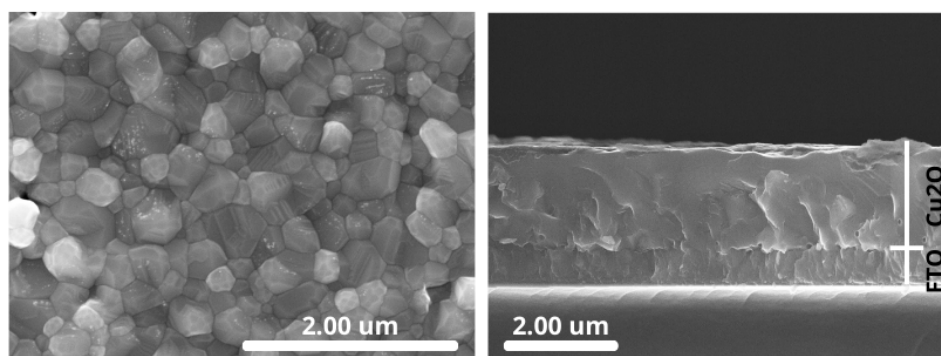


Figure 1: SEM images of annealed layers, left panel: Top view; right panel: cross-section

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@ntis.zcu.cz

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

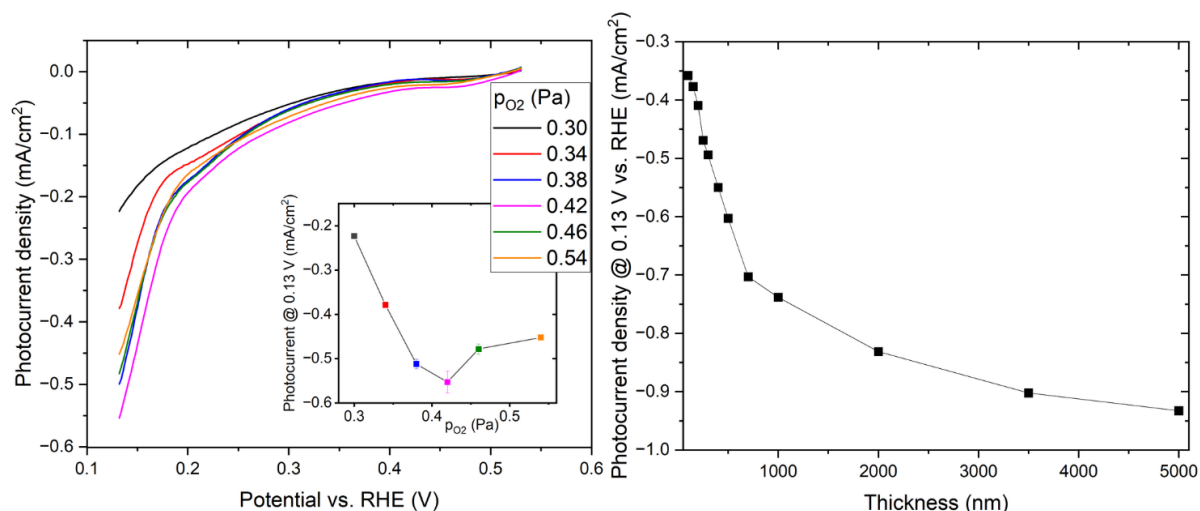


Figure 2: Left panel: The influence of oxygen partial pressure on PEC properties; right panel: layer thickness influence on maximal photocurrent density

2 Results

In this study, we utilize reactive HiPIMS to synthesize advanced Cu_2O materials for water splitting. A single phase composition was achieved after careful optimisation of oxygen partial pressure. Both, increased temperature of the substrate during deposition and post-annealing, led to higher crystallinity of the intended layer. As a result a fully densified single-phase layers with higher crystallinity were prepared, as shown in SEM images in Fig. 1.

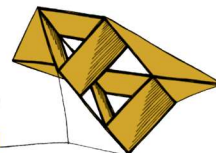
The effect of deposition parameters on the photoelectrochemical properties of the layers was also investigated. Crystallinity of the layers shows to be crucial as higher crystalline layers achieved higher photocurrent densities. As an optimal combination was identified substrate temperature of 500 °C and post-annealing in 600 °C. Optimal oxygen partial pressure of 0.42 Pa lead to the highest photocurrents (shown on left panel of Fig. 2) as the resulting layer do not include excessive ions of Cu or O. It was also found, that the higher the thickness of the layer, the higher the photocurrent densities, with saturation in the region of μm (right panel of Fig. 2). Fully optimised Cu_2O layer prepared by reactive HiPIMS in this work achieves photocurrent densities of -0,93 mA/cm² at 0,13 V vs. RHE, which rank among the highest reported for single-phase Cu_2O in the literature. That indicates promising opportunities for further advancements in multilayer systems and other PEC water splitting applications based on Cu_2O .

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.

References

- IEA (International Energy Agency), Global Hydrogen Review 2024, Paris (2024).
- Wick, R and Tilley, S.D., Photovoltaic and Photoelectrochemical Solar Energy Conversion with Cu_2O . *J. Phys. Chem. C*, 119 (2015) 26243-26257.



Synthesis and characterization of HiPIMS-deposited Zr–O–N films for potential water splitting applications

Minh Thanh Vu¹, David Kolenatý, Jiří Čapek, Stanislav Haviar, Radomír Čerstvý, Petr Zeman

1 Introduction

Zirconium oxynitride (Zr–O–N) films are renowned for their wear and corrosion resistance, temperature sensing capabilities, and optical properties [1,2]. Furthermore, these semiconductor materials demonstrate significant potential for water splitting applications, which are essential for sustainable hydrogen production. Recent studies have identified Zr_2ON_2 as one of the promising candidates for photoanodes due to its favorable band gap for efficient visible light absorption [3,4]. However, the synthesis of these oxynitrides remains underexplored, primarily due to the challenges associated with incorporating nitrogen into their structure.

In this work, we use the benefits of reactive high-power impulse magnetron sputtering (HiPIMS) to prepare Zr–O–N films across a wide composition range. Reactive HiPIMS is a progressive physical vapor deposition technique that achieves the significant ionization of sputtered species and the substantial dissociation and ionization of reactive gas species within the high-density discharge plasma near the target. This enables the synthesis of the Zr–O–N films with high crystallinity at low substrate temperatures while allowing precise tailoring oxygen and nitrogen content. Specifically, the films were deposited on Si and glass substrates at varying oxygen and nitrogen gas flow rates by sputtering a single Zr target. The substrates were unheated or ohmically heated to 500°C. Following depositions, the films were thoroughly analyzed by several analytical techniques such as wavelength-dispersive X-ray spectroscopy, X-ray diffraction, scanning electron microscopy and spectrophotometry.

The analyses of the elemental composition and the crystal structure revealed a continuous increase in nitrogen content, which is accompanied by the evolution of the structure from ZrO_2 through Zr_2ON_2 to ZrN as the nitrogen-to-oxygen ratio increases. This compositional and structural variation facilitates the tuning of material properties such as the electrical conductivity and bandgap width. These properties are essential for effective visible light absorption while maintaining proper alignment of the bandgap with respect to water splitting reactions. The bandgap energies were accurately determined using Tauc analysis demonstrating how variations in deposition parameters influence the electronic structure of the Zr–O–N films. The role of substrate temperature as well as the effects of post-annealing on the crystallization of the films will be also presented in detail.

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.

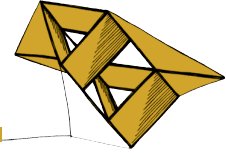
¹ student of the doctoral degree program Applied Sciences, field of study Plasma Physics, e-mail: Email: thanh@ntis.zcu.cz

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.

Sekce

Informatika, kybernetika



An Analytical Design Method for Power System Stabilizers Based on $\mathcal{H}_\infty$ Specifications

Michal Brabec¹, Lukáš Dostál²

1 Introduction

The power system is exposed to several fault conditions during its operation, such as short circuits, sudden load changes, or line disconnections. These disturbances create imbalances between mechanical and electrical quantities, leading to dangerous electromechanical oscillations, which manifest themselves as fluctuations in the angles between the rotors of synchronous generators. If these oscillations are not effectively damped, they can result in loss of synchronization and, in extreme cases, lead to the collapse of the power system. Standard control elements, such as Automatic Voltage Regulators (AVR), are often insufficient to dampen these oscillations on their own. For this reason, Power System Stabilizers (PSS) are commonly used in conjunction with AVR to ensure the dynamic stability of synchronous generators by actively damping electromechanical oscillations. The design of these stabilizers is a highly challenging task because of the dynamic nature of the controlled systems and the inherent uncertainty present in the system.

2 Identification and Modeling of the Key Dynamic Properties of the Controlled System

Knowledge of the key dynamic properties of synchronous generators is crucial for the design of stabilization controllers and the diagnosis of the behavior of the power system under various operating conditions. The identification of these properties can be effectively achieved using the frequency sweep method, which is based on harmonic excitation with a smoothly changing frequency. Building on this principle, an automatic identification method was proposed in Dostál and Schlegel (2023), capable of handling the inherent small non-linearities.

This approach enables the safe and relatively quick identification of a non-parametric frequency model, which can then be suitably approximated by a parametric model that serves as the foundation for the design of PSS parameters. By carefully selecting the excitation range and considering operational constraints, the risk of undesirable phenomena during the measurement process is minimized.

3 Analytical Design Method

The new approach to PSS design builds upon the previously published analytical method based on $\mathcal{H}_\infty$ specifications for the design of fixed-structure controllers with two tunable pa-

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

² student of the doctoral degree program Cybernetics, field of study Robust controllers with constrained structure, e-mail: lukasdos@ntis.zcu.cz

rameters described in Brabec and Schlegel (2023, 2024). It differs from conventional design approaches in that, rather than numerical optimization, it employs an analytical computational procedure for the transformation of various $\mathcal{H}_\infty$ design constraints to the boundary of the $\mathcal{H}_\infty$ region in the parametric plane of the controller. The $\mathcal{H}_\infty$ specification can also be extended with requirements for the damping of resonant modes (σ -stability, ξ -damping). Through the intersection of regions in the parametric plane, it is possible to implement both single and multiple model design and, at the same time, demand fulfillment of several constraints. The design procedure can be performed iteratively, which improves the resulting behavior of the controller and increases its order.

4 Application of the Analytical Design Method to PSS

$\mathcal{H}_\infty$ specifications are particularly well-suited for the design of PSS, where one of the standard requirements is to limit the maximum power gain within a specific frequency range. While publications Brabec and Schlegel (2023, 2024) primarily focus on designing affine structures (e.g., PI, PD, and PR controllers), the proposed analytical computational procedure is not restricted to this configuration. It can also be applied to alternative structures, such as the general Lead-Lag compensator, which is a key component of PSS structures. The design process is simply illustrated in Fig. 1.

The method is highly general and flexible, enabling the design of a wide range of standard PSS structures with washout filters and lead-lag compensators (as defined in *IEEE Std 421.5-2016*). Moreover, it is applicable not only to nominal models of the controlled system but also to models with non-structural uncertainty or model sets that cover different operating conditions.

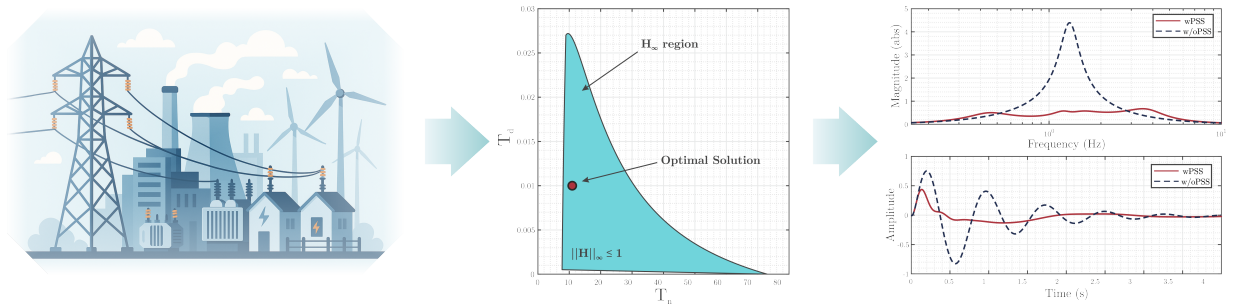


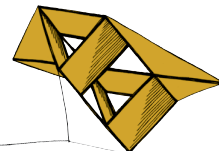
Figure 1: Illustration of the Power System Stabilizers design process: Models identification → PSS design in the parametric plane T_n - T_d of the Lead-Lag compensator → Verification of solutions with (—) and without (---) the PSS component based on active power characteristics.

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., 2024. Analytical Design of Controllers with Two Tunable Parameters Based on $\mathcal{H}_\infty$ Specifications for Dead-Time Systems, White Paper, www.pidlab.com.
- Dostálek, L. and Schlegel, M., 2023. The Frequency Response Estimation Using Swept-sine Excitation with Observer, 2023 24th International Conference on Process Control (PC), Strbske Pleso, Slovakia, pp. 72-77.



Monitorovací systém pro včelaře

Pavel Březina¹

1 Úvod

Včely jsou klíčovými opylovači a hrají zásadní roli v ekosystémech po celém světě. Péče o silná a zdravá včelstva vyžaduje pravidelný dohled, avšak tradiční kontroly úlů jsou časově i fyzicky náročné. Jejich omezená frekvence zvyšuje riziko, že dojde k přehlédnutí raných příznaků nemocí, rojení či ztráty matky. Moderní systémy pro real-time monitorování umožňují tyto události včas detekovat nebo dokonce předpovídat pomocí senzorů a zvukové analýzy.

Stávající komerčně dostupná řešení sice nabízejí širokou škálu monitorovacích funkcí, avšak jejich cena je často neúměrně vysoká a pro hobby včelaře finančně nedostupná. Navržený monitorovací systém přináší řešení za dostupnou cenu bez kompromisů na množství poskytovaných informací.

2 Monitorovací systém

Výsledný monitorovací systém se skládá z několika funkčních celků: sběru dat z úlu, jejich přenosu pomocí protokolu MQTT a následného ukládání a vizualizace na domácím serveru v reálném čase.

Pro měření uvnitř úlu byl vyvinut vlastní hardwarový modul, navržený s důrazem na solární napájení a modularitu. Zařízení je postaveno na mikropočítači ESP32-S2 Mini, pro který byl implementován specializovaný software zajišťující komunikaci se senzory a lokální vyhodnocení tří neuronových sítí přímo na zařízení zpracovávající zvukové záznamy.

Systém monitoruje parametry úlu jako je hmotnost, teplota na třech místech, vlhkost a výstupy tří neuronových sítí — konkrétně odhad síly včelstva, velikosti populace a stavu matky. Kromě toho také sleduje systémové veličiny jako napětí baterie, stav napájení, sílu bezdrátového signálu, teplotu mikropočítače a další. Naměřená data jsou bezdrátově odesílána pomocí protokolu ESP-NOW na přenášeč jednotku.

Software přenášeč jednotky je navržen s důrazem na minimalizaci spotřeby mobilních dat. Naměřená data jsou komprimována do úsporného binárního formátu a následně odesílána pomocí protokolu MQTT přes mobilní síť (méně než 1 MiB na úl za měsíc při toku 32 872 MQTT zpráv na úl – jedna zpráva obsahuje 23 monitorovacích hodnot).

Zprávy jsou přijímány domácím serverem, který zajišťuje jejich zpracování a ukládání do InfluxDB – databáze časových řad. Naměřené veličiny jsou následně vizualizovány v reálném čase prostřednictvím aplikace Grafana – specializovaný vizualizační software poskytující webové rozhraní (Obrázek 3).

¹ student navazujícího studijního programu Kybernetika a řídicí technika, obor Kybernetika, specializace Automatické řízení a robotika, e-mail: brezina1@students.zcu.cz

3 Výsledky

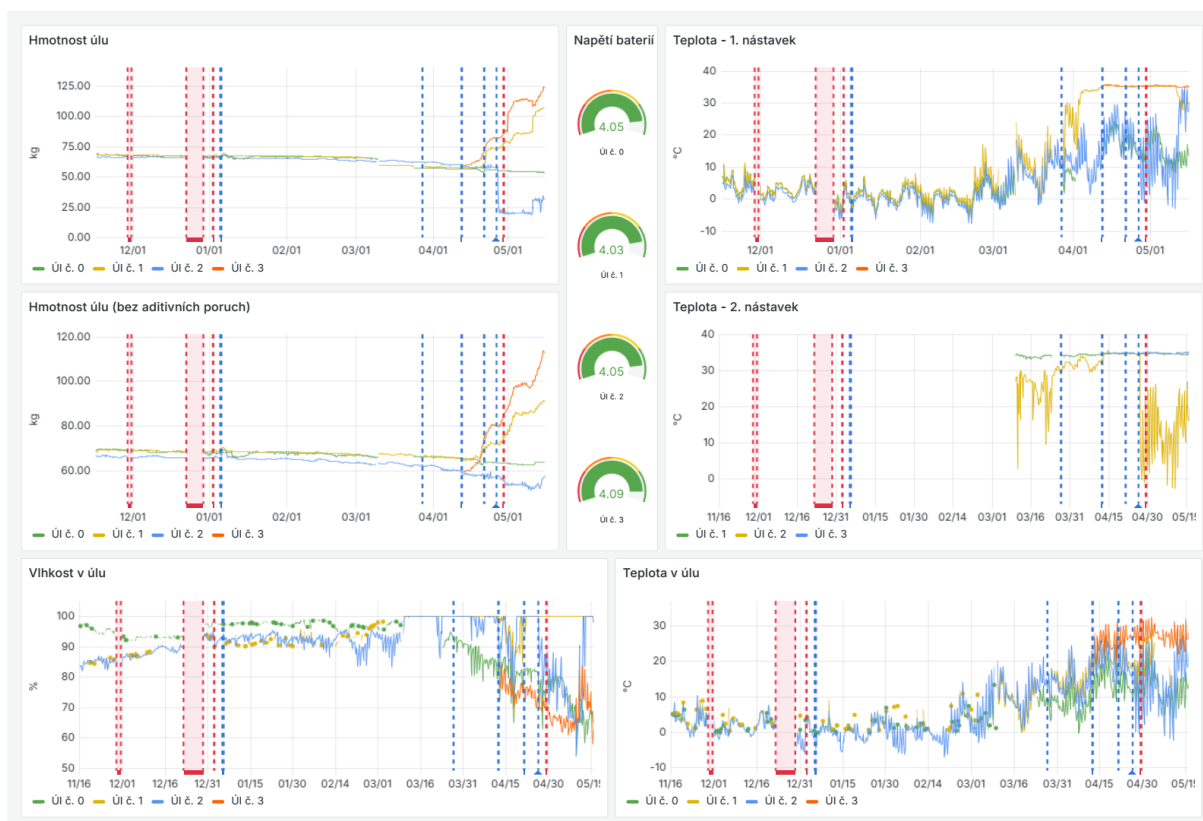
System byl nainstalován do čtyř úlů a běží téměř nepřetržitě od počátku nasazení (říjen 2024), což prokazuje jeho provozuschopnost. V průběhu implementace však bylo nutné provést několik neplánovaných úprav hardwaru, jejichž provedení se kvalitou značně liší od zbytku (Obrázek 2). Tyto improvizované změny zároveň dávají prostor pro budoucí optimalizaci návrhu a implementaci nových funkcí. Natrénované neuronové sítě poskytují slibné výsledky, které lze dále zpřesnit pomocí fine-tuningu přizpůsobeného konkrétnímu včelstvu a typu mikrofону.



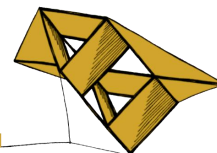
Obrázek 1: Sběrný modul měření přidělaný na záda úlu (19/10/2024).



Obrázek 2: Sběrný modul měření bez magnetického krytu (12/04/2025).



Obrázek 3: Ukázka webové vizualizace s vybranou podmnožinou naměřených hodnot. Červeně označené oblasti indikují výpadky systému, modře označené oblasti zachycují manipulaci s úly.



Zpracování obrazu pro automatické stříhání větví

Jakub Gaier¹

1 Úvod

V laboratoři výzkumného centra NTIS se vyvíjí vozík s robotickým ramenem, který bude zastříhovat větve vinné révy na vinicích. Automatizace takové činnosti vyžaduje počítačové vidění, které bude navigovat robotické rameno na správné místo řezu. Automatizace by usnadnila těžkou a monotónní práci vinařů několikrát do roka. Zdejší vývoj je zaměřen na řízení robotického ramena, ale v rámci jeho testování bylo třeba navrhnout jednoduchý prototyp pro zpracování obrazu, který bude fungovat v laboratorních podmínkách a nebude spoléhat na umělou inteligenci. Stříhání z pohledu robota vyžaduje znalost dvou parametrů, místa stříhání a orientaci nůžek, jež lze vyjádřit směrovým vektorem větve v místě větvení. Pro účely testování se nejprve vyzkoušelo rozpoznávat místa větvení keře, to se ale ukázalo jako příliš složité na rychlé prototypování, a proto se nakonec přešlo na hledání uzlin bambusu.

2 Popis experimentu a vývoje algoritmu

Vizuální data byla pořízena z hloubkové kamery Intel RealSense D435. Rozlišení kamery se nastavilo na nejvyšší možné, 1280x720 px, aby se zachytily křivky tenkých větví. Vysoké rozlišení omezuje velikost snímkové frekvence, která mohla být pouze 6 Hz, ale pro zdejší aplikaci není vysoká rychlost stěžejní, protože samotné zpracování obrazu je řádově v sekundách. Robot si tak bude muset vystačit s vizuální zpětnou vazbou, která je řádově pomalejší než smyčky regulátorů. Zbytek laditelných parametrů kamery byl nastaven na automatické zaměřování s automatickou expozicí. Kamera umožňuje snímat i hloubku obrazu. Pro snímání hloubky byl využit i infračervený projektor, který měl vylepšit kvalitu měření. Ovšem přesnost senzoru nedovolila měřit vzdálenost tenkých větví a reálně se dalo pracovat pouze se vzdáleností stonku. Algoritmus počítačového vidění tak musel být založen na zpracování barevného obrazu a hloubka sloužila jen jako pomocná informace pro následnou 3D lokalizaci.

První experimenty se prováděly na keři s hnědými větvemi umístěném před bílou zdí. Idea hledání bodů větvení spočívala v tom, že lze větve aproximovat po částech úsečkami. Úsečky lze pak poskládat do jakýchsi řetězců, které popisují původní větve. Stonek od větví odlišuje fakt, že se na něj napojuje největší počet větví; z pohledu komponent má nejvíce sousedů. Místa větvení se určili pomocí bližších konců sousedních řetězců vůči stonku a orientace by byla dána krajní úsečkou řetězce. Algoritmus pro zpracování obrazu se skládal z následujících kroků.

- Nejprve byl obraz předzpracován pomocí Cannyho hranového detektoru.
- Následně byly z hran odhadnuté úsečky pomocí Houghovy transformace.
- Úsečky byly rozděleny na horizontální a vertikální a jednotně orientovány.

¹ student navazujícího studijního programu Aplikované vědy a informatika, obor Kybernetika a řídicí technika, specializace Automatické řízení a robotika, e-mail: gaier@students.zcu.cz

- Orientované úsečky se podle pravidel spojovaly v řetězy.
- Řetězy se poté opakovaně spojovaly navzájem, až nakonec tvořily celé větve. Finální podoba řetězů je zobrazena na obr. 1.
- Z větví byl vybrán stonek a podle něj se určila místa větvení.

V průběhu řešení jsem se potýkal s problémy, jak spojovat úsečky do řetězů, aby byly splněny podmínky, že se řetěz může rozvíjet pouze na koncích. Musel se tedy rozlišovat konec a začátek. Dále byl problém s tím, že úsečky mohou být orientované libovolně, tudíž i vertikálně, a to znemožňuje použít popis pomocí lineárních funkcí. Dále samotný proces stavění řetězů vyžaduje kontrolování za jakých podmínek se budou sousední řetězy spojovat a za jaké konce se spojí. Ani fakt, že větve mají určitou tloušťku a tudíž se produkovali úsečky na obou hranách větví, nepřispíval k jednoduchosti řešení. Zároveň byla brána v potaz dobu výpočtu, takže stavění algoritmu vyžadovalo vymýšlení efektivních pravidel. Nakonec z tohoto experimentu vzešel program, který byl schopen nalézt místa větvení, ale měl problém s konzistentností a výpočetní náročností. Navíc pravidla algoritmu byla tak komplexní, že se program velmi těžko ladil a jeho reálné použití na robotovi by nejspíše nevedlo k robustní navigaci.

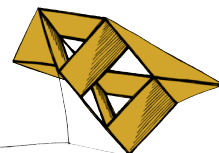


Obrázek 1: Popis větví keře pomocí řetězů orientovaných vektorů

Rozhodlo se proto od tohoto postupu upustit a pro zájmy testování robota se přešlo na druhý experiment s bambusem. Bambus je tvarově mnohem jednodušší než větve keře a lze na něm rozpoznávat jeho uzliny. Algoritmus pro jejich nalezení spočívá v hledání jasových změn červené složky obrazu v naprahaných podlouhlých oblastech. Program je i rychlý natolik, že umožňuje zpracovávat video záznam a byl obohacen o kontrolu konzistence detekcí uzlin v průběhu nahrávky. Na finální verzi programu jsem se již nepodílel, ale nakonec se povedlo vytvořit prototyp počítačového vidění vhodného pro laboratorní účely.

Poděkování

Tato práce byla podpořena projektem SGS-2025-20.



Detekce zrcadlové symetrie v 3D geometrických modelech proměnných v čase

Vít Gregor¹

1 Úvod

Symetrie je důležitou vlastností geometrických objektů. Jedná se o geometrickou transformaci, vůči které je objekt invariantní, tedy aplikací této transformace nedojde k jeho změně. Cílem této práce bylo navrhnout, implementovat a vyzkoušet metodu pro detekci symetrie v objektech měnících se v čase.

2 Metody pro detekci symetrie

Při studiu metod pro detekci symetrie bylo zjištěno, že nejsou známy metody, které by se přímo zabývaly detekcí symetrie v několika snímcích animace se zajištěním koherence, tedy plynulé návaznosti výsledných rovin sousedních snímků na sebe. Proto byly v rámci této práce analyzovány pouze metody zpracovávající statické objekty, a následně bylo navrženo, jakým způsobem je možno tyto metody použít pro detekci symetrie v animacích.

Pro vývoj metod na zlepšení koherence byla jako základ vybrána metoda založená na diferencovatelné míře symetrie popsaná v Hrudá et al. (2022). Tato metoda pro každou kandidátní rovinu počítá míru symetrie. Na základě míry symetrie se vybere pět nejlepších kandidátních rovin. Z těchto pěti rovin se pak vybere jedna nejlepší. Tohoto rozdělení na dva kroky při výběru roviny bylo v rámci této práce využito ke zlepšení koherence, kdy po vybrání pěti nejlepších kandidátních rovin následuje modifikace jejich hodnot míry symetrie, která bere v úvahu to, jaké výsledné roviny symetrie byly získány v předchozích snímcích.

3 Řešení

Jednotlivé metody pro hledání symetrie v objektu měnícím se v čase jsou navrženy podle tří scénářů. První scénář je založen na tom, že se každý snímek zpracuje Hrudovou metodou, a koherence se nezajišťuje. Tato metoda slouží především pro porovnání s následujícími metodami, jejichž cílem je koherenci mezi jednotlivými snímky zlepšit.

Druhý scénář je založen na upravování měr symetrie v programu Lukáše Hrudy ve fázi, kdy metoda již vybrala pět nejlepších kandidátních rovin z hlediska míry symetrie. V této fázi lze míru symetrie upravit pomocí vah, kdy bude přiřazena větší váha kandidátním rovinám podobným těm, které se vyskytly v nedávných výsledcích. Další možností je nastavení nekonečné míry symetrie rovině podobné té, kterou chceme udržet.

Třetí scénář je založen na tom, že není nutné Hrudovu metodu nasazovat v každém snímku, ale pouze v klíčových snímcích, které se určují buď uniformně, nebo na základě překročení maximální chyby symetrie počítané podle rovnice 1, kde $\mathbf{n}$ je normálový vektor

¹ student studijního programu Informatika a její specializace, specializace Počítačová grafika, e-mail: gregorv@students.zcu.cz

roviny symetrie a $\mathbf{x}_i$ představuje bod objektu, kterých je celkem n .

$$\epsilon = \left| \frac{1}{n} \sum_{i=1}^n \frac{\mathbf{n}^T \mathbf{x}_i + d}{\sqrt{\mathbf{n}^T \mathbf{n}}} \right| \quad (1)$$

Jedna z metod třetího scénáře pracuje tak, že v klíčovém snímku rozdělí objekt na dvě části, a poté sleduje, jak se pohybují centroidy těchto částí v čase a prokládá mezi nimi rovinu. Druhá metoda je založena na lineární interpolaci mezi dvěma klíčovými snímky.

4 Výsledky

Výsledky se hodnotily především z pohledu koherence, tedy jak plynule se výsledná rovina v průběhu animace mění. Výsledky dopadly nejhůře pro první scénář, u kterého nebyla koherence zajišťována.

Výrazně lépe dopadly výsledky pro druhý scénář, kdy bylo detekováno podstatně méně neočekávaných rovin symetrie, tedy takových rovin, které byly výrazně či zcela odlišné od výsledných rovin sousedních snímků.

Nejlepší výsledky z pohledu koherence byly získány pro třetí scénář, především pro metodu interpolace. V některých případech ale byly nalezeny roviny symetrie, které byly posunuté či vychýlené z pohledu toho, jak symetrii vnímají uživatelé. Ukázka výsledku pro metodu lineární interpolace je vidět na obrázku 1. Metody navržené v rámci třetího scénáře fungovaly podstatně rychleji než u obou předchozích scénářů.



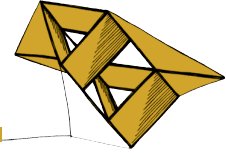
Obrázek 1: Ukázka výsledku pro snímky 1 a 90 animace postavy dělající stojku při využití lineární interpolace.

5 Závěr

V rámci této práce byly navrženy tři scénáře pro detekci symetrie v objektech měnících se v čase. Bylo zjištěno že nejlépe funguje z hlediska koherence nasazení Hrudovy metody pouze na klíčové snímky a dopočtení ostatních interpolací, popřípadě pomocí centroidů. Kromě toho byly vyvinuty metody pracující s mírami symetrie rovin přímo uvnitř Hrudova programu, jejichž hlavní nevýhodou byla dlouhá doba provádění.

Literatura

Hruda, L., Kolingerová I. a Váša, L. (2022). *Robust, fast and flexible symmetry plane detection based on differentiable symmetry measure*. The Visual Computer. Dostupné z: 38.10.1007/s00371-020-02034-w.



Measurement Dimensionality and Its Impact on Position Error Covariance in Tracking Applications

Michael Gulyj¹

1 Introduction

Precise localization of aircraft is essential for air traffic control, surveillance, and operational safety. In tracking systems, this must be achieved without any onboard sensors. The flying object emits a signal that can be captured by a network of ground-based receivers. Since the exact time of emission is unknown, the system relies on Time Differences of Arrival (TDOA) measurements. If the sensors are time synchronized, the TDOA measurements contain enough information to accurately determine the position of the aircraft.

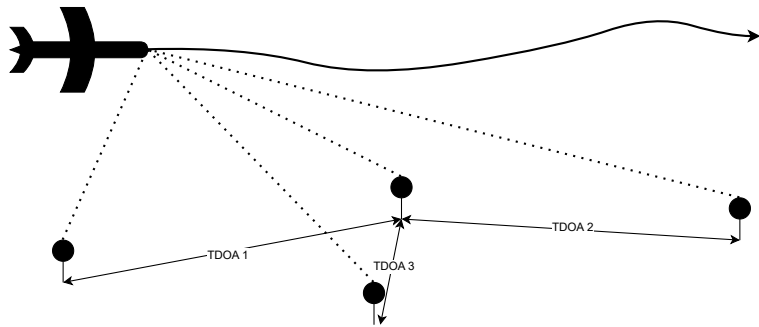


Figure 1: Illustration of TDOA-based localization

2 The goal of the paper

The position of the aircraft cannot be determined exactly since the measurements are subject to uncertainties. The localization algorithms thus not only provide the position estimate but also quantify its uncertainty using an error covariance matrix. The goal of this work is to examine how uncertainty develops along various trajectories, especially when one or more sensors fail or become unavailable.

3 Localization approaches

Two different localization approaches are compared in this work. The first directly processes the TDOA measurements and uses a nonlinear filter, specifically the unscented Kalman filter, to provide the estimates. The second approach transforms the TDOA measurements into virtual Cartesian position measurements and then applies the standard Kalman filter. This transformation is done by computing the intersection of a hyperboloid and 2 planes, a method known as hyperbolic positioning Al-Samahi et al. (2020). A key challenge of this method is that the virtual measurement has to be three-dimensional, regardless of the number of available sensors. When a measurement is missing, the algorithm must compensate for it by using state prediction, and the uncertainty along the corresponding directions is implicitly modeled as infinite.

¹ student of the master's degree program Cybernetics and Control Engineering, e-mail: gulyjm@students.zcu.cz

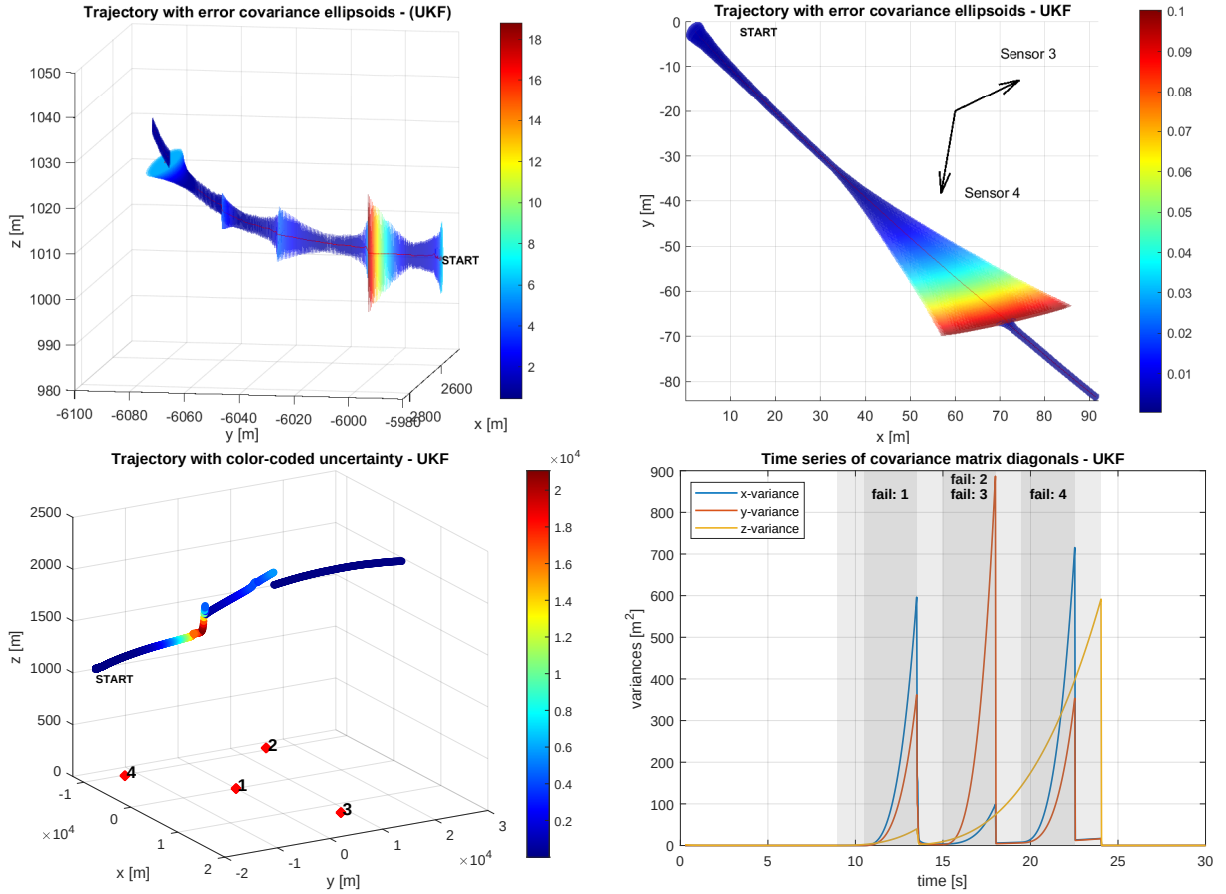


Figure 2: Examples of graphical visualizations generated by the developed framework

4 Key findings

Based on simulations¹, the following key conclusions can be drawn. (i) When all sensors are operational, or when a single sensor failure lasts less than ten seconds, both approaches produce nearly identical outputs. (ii) In more complicated scenarios, the approach with nonlinear measurements gives better results in terms of consistency and estimation errors. (iii) The uncertainty tends to increase in the direction of the missing sensor. (iv) The flight direction does not have a direct impact on the error covariance matrix. (v) The most significant factor influencing the estimation quality is the geometry, i.e., the aircraft's position relative to the sensor layout.

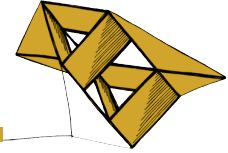
Acknowledgement

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

References

- Al-Samahi, S.S.A., Zhang, Y., and Ho, K.C. (2020) Elliptic and hyperbolic localizations using minimum measurement solutions. *Signal Processing*, **167**, 107273. Elsevier.
- Siřiste, M. (2024) Quality assessment of manoeuvring object tracking by radar multilateration. *Diploma thesis*, Západočeská univerzita v Plzni.

¹MATLAB simulation framework can be found at https://github.com/MichaelGulyj/DP_Gulyj. It was developed as an extension to the diploma thesis Siřiste (2024)



Do Compact Multilingual Large Language Models Speak Czech?

Jakub Hejman¹

1 Introduction

The recent popularity explosion of Large Language Models (LLMs) drives a similar growth in their demand by industry practitioners. However, the widely available API offerings are often not suitable for proprietary use cases. This work investigates the Czech capabilities of current LLMs in a size range that is practical for self-hosting.

In the previous paradigm of pretrained transformers such as BERT or T5, there was a need to pretrain language-specific models due to the large performance gap. With the scale-up and fast development of decoder-only pretraining approaches, have larger and more general LLMs reached the effectiveness of specialized models for narrow tasks?

2 Approach

Pretrained LLMs are evaluated on two traditional NLP tasks—Sentiment Classification and Coreference Resolution—in the Czech language to gain a wider perspective. The models are finetuned using QLoRA with a context length of up to 4096.

2.1 Sentiment Classification on the ČSFD Dataset

The dataset selected for Sentiment Classification experiments is originally a collection of scraped ČSFD reviews. Review texts are used as model inputs and ratings were converted from the number of stars in the original system to polarity labels. The simple prompt for this task is described in Listing 1.

We choose a label mapping where the prefix of each completion is different, aiming to avoid extended generations from the model, in favor of sampling a single token.

```
Byli sme na tom s kamosema v kine. A jedina vyhoda byla ta, ze nas tam bylo  
akorat tolik aby se to promitalo a clovek se tam mohl jak chtel rozvalit  
a bavit nahlas  
### HODNOCENÍ: Nelíbí se mi.
```

Listing 1: An instantiated prompt for Sentiment Classification. The task input is highlighted in yellow, the prompt separator in green, and the expected generation in blue. The possible completions are “Líbí se mi” (“Like it”) for positive, “Bylo to okay” (“It was okay”) for neutral, and “Nelíbí se mi” (“Dislike it”) for negative.

¹ student of the master degree program Informatics and Its Specializations, field of study Natural Language Processing, e-mail: hejmanj@students.zcu.cz

2.2 Coreference Resolution using CorefUD

The task of Coreference Resolution involves tagging entity mentions and creating mention clusters corresponding to each entity. This is still an active area of research with competitions such as CRAC happening yearly. CorefUD 1.1, the dataset used in the 2023 round of the competition, along with baseline from that round are used for comparison.

In the generation-oriented approach, a mention annotation is added after every word token. Annotations may consist of multiple tags to handle nesting, each denoting an entity ID along with parentheses to distinguish spans that are being opened, closed or consist of only the single word. This custom format is shown in Listing 2.

As the official evaluator requires valid inputs, a post-processing step is necessary to normalize LLM outputs. The outputs are aligned using an edit-distance-like algorithm and mismatched opening or closing tags are resolved by conversion to single word tags.

Váš obecně platný dotaz je připraven zodpovědět spolupracovník Profitu .
 V případě , že se nedovoláte , svůj dotaz namluvte na telefonní záznamník .

LABEL

Váš|(e1) ,(e2 obecně platný dotaz|e2) je připraven zodpovědět
 spolupracovník|(e3 Profitu|e3) . V případě , že se nedovoláte
 , svůj|(e2) ,(e1) dotaz|e2)) namluvte na telefonní záznamník .

Listing 2: Coreference Resolution example. The task input is yellow, and the separator green. The expected output is highlighted according to the entities.

3 Experimental Results

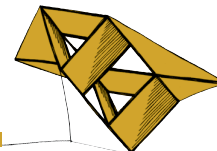
Model	ČSFD (acc)	CorefUD (cs score)
FERNET - Šmíd, J. et. al. (2023)	88.2	–
CRAC baseline - Žabokrtský, Z. et al. (2023)	–	65.22
CohereLabs/aya-expanse-8b	90.1	57.25
meta-llama/Llama-3.1-8B	84.2	61.02

Table 1: Experimental result.- Our results on CorefUD use the public dev set.

The main results in Table 1 demonstrate that LLMs are capable of competing with specialized models; however there is still more work to be done before they become the go-to for industry applications. We note that while the quality in Czech lags behind the quality in English, the gap is somewhat smaller than using past approaches.

References

- Šmíd, J. and Přibáň, P. (2023) ‘Prompt-Based Approach for Czech Sentiment Analysis’ *Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing*. Varna, Bulgaria: INCOMA Ltd., Shoumen, Bulgaria, pp. 1110–1120.
- Žabokrtský, Z. et al. (2023) ‘Findings of the Second Shared Task on Multilingual Coreference Resolution’ *Proceedings of the CRAC 2023 Shared Task on Multilingual Coreference Resolution*. Singapore: Association for Computational Linguistics, pp. 1–18.



Autoregressive Upscaling of Sparse Single-Cell Data Improves Interpretability

Tomáš Honzík¹, Lukáš Kuhajda², Daniel Georgiev³

1 Introduction

scRNA-seq is a promising data modality in medical research, offering relatively high-resolution information about individual cells [1]. However, it suffers several limitations, for example high sparsity and significant noise in the non-zero values [2]. As a result, using raw reads for downstream analysis—such as fine-grained cell typing is fundamentally constrained [3]. In a sense, much of the useful information is hidden behind zeros. We can ask: if we had sequenced 100 times more RNA molecules, would a gene still read as 0? Or would it appear as 1, 10, or even 50? This study presents a deep learning model designed to answer exactly that: to upscale the measured values and evaluate whether the upscaled data improves downstream performance on supervised type classification. The training dataset included PBMCs from blood and bone marrow, totaling 14 million cells. The data were downloaded from <https://cellxgene.cziscience.com/> [4].

2 Methods

The model backbone is a simple multilayer perceptron (MLP), enhanced with practical techniques such as layer normalization and skip connections [5-6]. The novelty of our method does not lie in the architecture itself, but in how the inputs and outputs are defined during training and inference. We begin with an intuitive explanation. Each cell is conceptualized as a bag of RNA tags, where each tag corresponds to one of 16,000 human genes. During training, we take one such bag and—without initially revealing any content—ask the model: "If we were to draw a random RNA tag from this bag, which gene do you think it would correspond to?" The model makes a prediction, which is then compared to the actual identity of the randomly drawn tag using categorical cross-entropy. This tag is then revealed to the model and added to the input. The process repeats. Once the bag is emptied and all tags are revealed, a new cell is selected and the procedure restarts. Through this iterative refinement, the model first learns a prior gene distribution, then progressively sharpens its predictions as more RNA context becomes available. Importantly, during inference, the model can generate RNA tags autoregressively—without access to actual measurements—modeling the joint distribution of gene co-expression. The actual training procedure is mathematically equivalent but optimized for numerical stability and efficiency. Specifically, we split each RNA count vector into two parts using a random ratio from Uniform(0,1). The input vector is transformed as $\log(x+1)$, while the output vector is normalized as $x/\text{sum}(x)$ to represent a probability distribution—reflecting the likelihood of drawing each gene tag. During inference, we can start either with an empty input (generating cells from scratch) or with a measured RNA vector (upsampling). We then autoregressively sample from the predicted distributions, always feeding the sampled tag back into the input.

¹ University of West Bohemia, Faculty of Applied Sciences, e-mail: thonzik@ntis.zcu.cz

² University of West Bohemia, Faculty of Applied Sciences, e-mail: kuhajdal@ntis.zcu.cz

³ University of West Bohemia, Faculty of Applied Sciences, e-mail: georgiev@kky.zcu.cz

3 Results

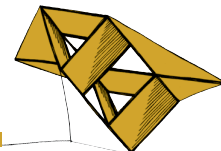
When the measurements from the dataset are upscaled, the natural question arises: “Does the upscaled version actually contain more information about the real physical cell, or is it just a fancy hallucination?” Here, we describe our methodology to evaluate this. First, we compared t-SNE plots of raw and upscaled measurements. Visual inspection consistently showed that raw t-SNE plots clearly separate major PBMC types such as T cells, B cells, and monocytes. However, separation at a higher granularity rarely went beyond distinguishing CD4 and CD8 T cells—and even that often failed. In Figure 3, we demonstrate the remarkable difference between raw measured cells and their upscaled counterparts for a subpopulation of NK and T-cells from a single donor (6,000 cells, each with 1k–20k RNAs). Note that all available information was used for the first graph, so the model genuinely imputes information that was not originally measured. To go beyond t-SNE plots (whose hyperparameters and normalization were selected fairly to avoid bias), we trained classifiers to predict 95 cell types. The classifier trained on upscaled data consistently achieved a lower error rate (21% vs. 27%). However, the absolute difference might seem smaller than one would expect from Figure 3. Interestingly, we observed a much larger improvement in loss, indicating that the predicted probabilities are more closely aligned with the true uncertainty. To illustrate this, we discarded the bottom half of least confident predictions in both models and found a stark difference: 1% error on upscaled data vs. 19% on raw data. Finally, we compared real cells from the dataset with cells generated entirely from scratch by the model. By mixing synthetic and real cells in a t-SNE plot, we provide weak but encouraging evidence that the generated cells plausibly follow the same underlying distribution.

Acknowledgement

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

References

- [1] Kolodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C., Teichmann, S.A. (2015) The technology and biology of single-cell RNA sequencing. *Molecular Cell*, 58(4), pp. 610–620. <https://doi.org/10.1016/j.molcel.2015.04.005>
- [2] Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D.J., Hicks, S.C., Robinson, M.D., et al. (2020) Eleven grand challenges in single-cell data science. *Genome Biology*, 21(1), p. 31. <https://doi.org/10.1186/s13059-020-1926-6>
- [3] Luecken, M.D., and Theis, F.J. (2019) Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology*, 15(6), e8746. <https://doi.org/10.15252/msb.20188746>
- [4] CZI Single-Cell Biology, et al. (2023) CZ CELLxGENE Discover: A single-cell data platform for scalable exploration, analysis and modeling of aggregated data. *bioRxiv*. <https://doi.org/10.1101/2023.10.30.563174>
- [5] Ba, J.L., Kiros, J.R., Hinton, G.E. (2016) *Layer Normalization*. Available from: <https://arxiv.org/abs/1607.06450>
- [6] He, K., Zhang, X., Ren, S., Sun, J. (2016) Deep residual learning for image recognition. *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778. Available from: <https://arxiv.org/abs/1512.03385>
- [7] van der Maaten, L., and Hinton, G. (2008) Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9(Nov), pp. 2579–2605. Available from: <http://www.jmlr.org/papers/v9/vandermaaten08a.html>



Modul pro statistickou analýzu medicínských dat v systému Czech Salivary Gland Database

Bc. Vojtěch Jelínek¹

1 Úvod

Nádory slinných žláz představují specifickou výzvu v oblasti onkologie hlavy a krku, a to jak z hlediska diagnostiky, tak i následné léčby. Jejich výskyt je relativně vzácný a jejich značná histologická heterogenita komplikuje jednotné léčebné postupy. Z těchto důvodů je nezbytné mít k dispozici nástroje, které umožňují systematický sběr, správu a statistickou analýzu dat o pacientech s těmito nádory.

Czech Salivary Gland Database (CSGDB) je specializovaný informační systém vyvinutý ve spolupráci s Fakultní nemocnicí Motol, jehož cílem je podpora klinického i výzkumného zpracování dat pacientů s nádory slinných žláz. Systém umožňuje evidenci pacientů, správu klinických studií a základní analýzu dat. Statistické výstupy hrají klíčovou roli při vyhodnocování léčebných výsledků, identifikaci rizikových faktorů i při klinickém rozhodování. Důraz je kladen na možnost agregace dat z více pracovišť, čímž se překonávají omezení plynoucí z malého počtu případů na jednotlivých klinikách.

2 Popis aplikace

Aplikace CSGDB je multiplatformní desktopová aplikace postavena na JavaScript technologii Electron, využívající knihovnu ReactJS pro tvorbu moderního uživatelského rozhraní. Architektura systému je třívrstvá – prezentační vrstva slouží k zadávání vstupních parametrů a zobrazování výstupů, aplikační logika na backendu zajišťuje zpracování požadavků a interakci s relační databází SQLite.

Uživatelské rozhraní bylo navrženo ve spolupráci s lékaři, což se promítá do podpory běžných klinických scénářů, jako je filtrování pacientů podle diagnózy, věku nebo typu výkonu. Dále je možné provádět pokročilé statistické analýzy, které zahrnují deskriptivní statistiku, inferenční testy a vizualizaci dat. Systém rovněž podporuje export výsledků do formátů PDF a Excel pro další zpracování.

3 Data sbíraná v aplikaci

Systém podporuje sběr širokého spektra klinických a demografických údajů. Zaznamenává se například věk a pohlaví pacienta, histologický typ nádoru, klinické stadium onemocnění, provedené chirurgické výkony, informace o recidivách, komplikacích a výsledcích dlouhodobého sledování. Identifikace pacientů je řešena pseudonymizací, aby bylo zajištěno dodržení požadavků GDPR.

Data jsou ukládána v normalizované databázové struktuře s podporou AES-256 šifrování citlivých údajů. Systém rovněž zahrnuje mechanismy pro validaci vstupních hodnot a jejich

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

standardizaci, což je klíčové pro následné statistické zpracování.

4 Cíl zadání

Hlavním cílem vývoje statistického modulu bylo vytvořit nástroj, který lékařům a výzkumníkům umožní snadno a rychle analyzovat dostupná klinická data. Modul byl navržen tak, aby umožňoval identifikaci vztahů mezi demografickými proměnnými, histologickým typem nádoru a léčebnými výsledky. Systém má rovněž sloužit k ověřování klinických hypotéz na základě statistických testů a tím přispět k personalizaci léčby.

Díky integraci modulu přímo do systému CSGDB odpadá nutnost externího zpracování, čímž se výrazně snižuje čas potřebný pro získání výsledků.

5 Implementace

Statistické výpočty jsou realizovány pomocí knihovny `jstat`, která poskytuje funkce pro základní deskriptivní statistiku (např. průměr, medián, směrodatná odchylka) i pro běžně používané inferenční testy. Modul aktuálně podporuje chi-kvadrát test, Fisherův exaktní test, Studentův t-test a neparametrický Mann-Whitneyův U test.

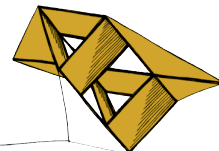
Při použití chi-kvadrát testu má uživatel možnost flexibilně definovat vlastní kategorizaci dat pro tvorbu kontingenční tabulky, včetně slučování kategorií. Tato funkce byla přidána na základě zpětné vazby od klinických uživatelů a umožňuje lépe reflektovat reálné potřeby datové analýzy. Stejný přístup je možné uplatnit i při použití Fisherova exaktního testu, což rozšiřuje využitelnost modulu i pro situace s malými četnostmi, kde je právě Fisherův test vhodnější než chi-kvadrát test.

Výsledky analýz jsou zobrazeny přehledně ve formě interaktivních tabulek a grafů. Pro potřeby klinické dokumentace je k dispozici export výsledků do formátů PDF a Excel. Modul byl testován na reálných datech za spolupráce s lékaři z FN Motol, přičemž důraz byl kladen nejen na přesnost výpočtů, ale i na uživatelskou přívětivost a intuitivnost ovládání.

6 Závěr

Statistický modul integrovaný do systému CSGDB představuje efektivní nástroj pro podporu klinického rozhodování a retrospektivní analýzy. Umožňuje provádění standardizovaných analýz přímo nad pseudonymizovanými daty bez nutnosti exportu do jiných nástrojů. To vede ke zkrácení doby potřebné pro zpracování výsledků, zvýšení konzistence výstupů a lepší reprodukovatelnosti závěrů.

V klinické praxi se modul již uplatnil při přípravě onkologických konzilií a zpracování retrospektivních studií. Do budoucna se plánuje rozšíření funkcionality o pokročilé statistické metody, multivariační modely a metody strojového učení, které umožní například predikci rizika recidivy či odpovědi na léčbu. Tím se posiluje potenciál systému CSGDB jako nástroje pro výzkum vzácných nádorových onemocnění a rozvoj personalizované medicíny.



Burzovní simulátor pro trh řízený limitními objednávkami

Vladimíra Kimlová¹

1 Abstrakt

Rychlost, anonymita a komplexní interakce účastníků činí z finančních trhů extrémně náročné prostředí pro analýzu i regulaci. Tato práce představuje simulační platformu pro modelování trhu řízeného limitními objednávkami, s důrazem na realističnost, modulární rozšiřitelnost a systematické testování obchodních strategií. Základem systému je realistický párovací algoritmus a implementace obchodních agentů včetně modelu manipulativního chování. Architektura simulátoru je optimalizována na nízkou latenci, vysokou propustnost a flexibilní konfiguraci tržních podmínek. Experimentální simulace potvrzují schopnost systému věrně reprodukovat dynamiku trhu a odhalovat dopady manipulace na jeho stabilitu a efektivitu. Výsledné řešení poskytuje efektivní nástroj pro výzkum mikrostruktury finančních trhů, algoritmického obchodování a detekce podvodných praktik.

2 Úvod

Cílem této diplomové práce je návrh a implementace **burzovního simulátoru** pro trh řízený limitními objednávkami, který umožňuje simulaci obchodování prostřednictvím autonomních i manuálních obchodních agentů a poskytuje nástroje pro analýzu a vizualizaci výsledků.

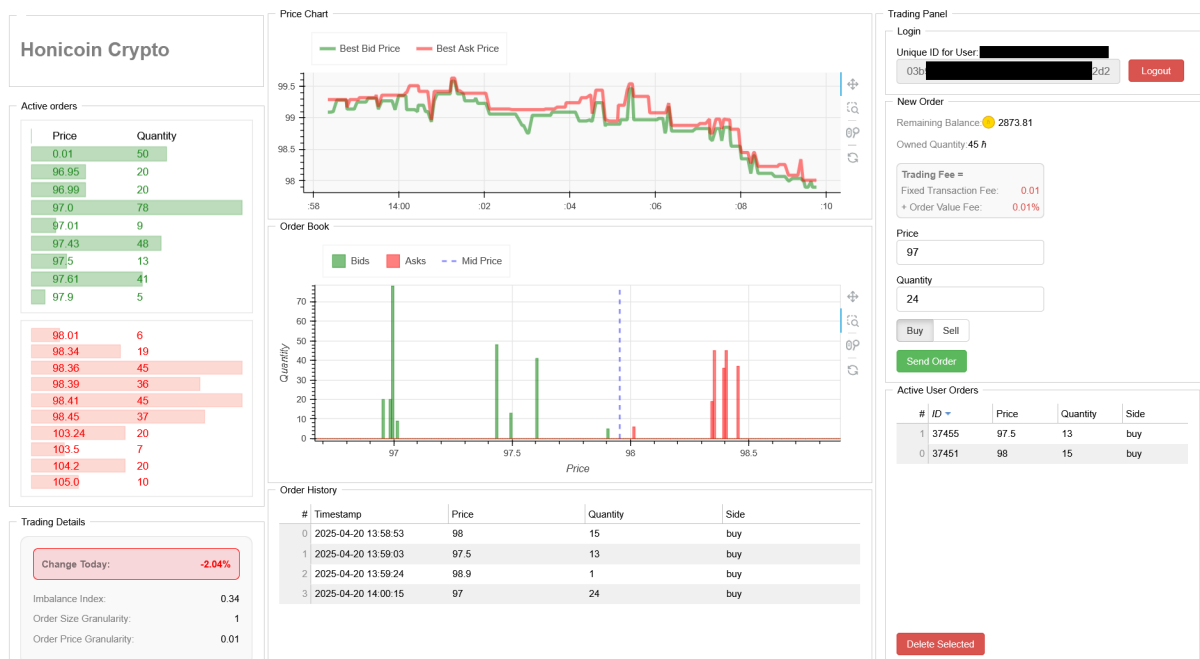
3 Východiska, analytická část

Limitní kniha objednávek (LOB) je základní strukturou elektronických trhů, která organizuje příkazy k nákupu a prodeji podle ceny a času (viz Abergel et al. (2016)). Párování probíhá zpravidla algoritmem **price-time priority**. LOB je dynamický systém ovlivňovaný interakcí tisíců účastníků. V kontextu simulace trhu je důležité modelovat nejen samotné mechaniky párování objednávek, ale i chování účastníků. K tomu slouží **agentové modelování**, které reprezentuje účastníky jako autonomní agenty s různými strategiemi a úrovní inteligence. Dále je důležité zohlednit výskyt manipulativních praktik, jako je **spoofing** (zadávání a rušení falešných objednávek s cílem ovlivnit tržní sentiment). Tyto jevy mají zásadní dopad na stabilitu a důvěryhodnost tržního prostředí, a proto jsou důležitým prvkem i v simulačním výzkumu.

4 Implementované řešení

Simulátor byl navržen jako **modulární a rozšiřitelný systém** umožňující realistické obchodování na trhu řízeném limitními objednávkami. Jádrem je **serverová aplikace** v jazyce Python, postavená na **asynchronní architektuře** s důrazem na nízkou latenci a vysokou propustnost. Obsahuje implementaci párovacího algoritmu podle principu **price-time priority** a správu uživatelských účtů i financí. Komunikace mezi komponentami využívá upravený pro-

¹ studentka navazujícího studijního programu Informatika a její specializace, specializace Zpracování přirozeného jazyka, e-mail: jivl@students.zcu.cz



Obrázek 1: Navržené webové uživatelské rozhraní pro manuální obchodování

tokol FIX (viz FIX Trading Community (2017–2025)), který je standardem ve finančním sektoru. Klientská část zahrnuje **webové rozhraní** pro manuální obchodování (viz Obr. 1) a API pro vývoj autonomních agentů. Byli implementováni reprezentativní obchodní agenti s různými strategiemi včetně jednoho manipulativního (**využívajícího spoofing**), přičemž agenti mohou být snadno konfigurováni či rozšiřováni. Systém dále obsahuje nástroje pro **sběr dat**, generování **statistických reportů** a **vizualizaci výsledků simulací**, čímž umožňuje systematické porovnání strategií a analýzu dynamiky trhu v různých podmínkách.

5 Dosažené výsledky

Simulátor byl otestován ve třech scénářích: zátěžový test pro ověření výkonnosti serveru, uživatelská simulace v podobě manuální burzovní hry a série simulací s autonomními agenty. Výsledky potvrdily stabilitu systému, korektní párování objednávek a schopnost reprodukovat charakteristické jevy tržní mikrostruktury včetně dopadů manipulativních strategií.

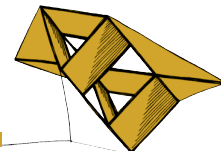
6 Závěr

Byl navržen a realizován robustní simulátor trhu řízeného limitními objednávkami s podporou autonomních a manipulativních agentů. Systém umožňuje testování a analýzu obchodních strategií v realistickém prostředí a slouží jako nástroj pro výzkum tržní dynamiky, algoritmického obchodování a detekci tržních anomálií.

Literatura

Abergel, F., Anane, M., Chakraborti, A., Jedidi, A. a Toke, I.M. (2016). *Limit Order Books*. Cambridge University Press. DOI: 10.1017/CBO9781316683040.

FIX Trading Community (2017–2025). *FIX Trading Community*. Dostupné z: <https://www.fixtrading.org>.



Software for Neurorehabilitation

Tomáš Kment¹

1 Introduction

Neurological disorders such as stroke, traumatic brain injury (TBI), spinal cord injury (SCI), and neurodegenerative diseases affect hundreds of millions worldwide. Feigin et al. (2025) representing the World Stroke Organization (WSO) reports that ≈ 93.8 million people lived with stroke in 2021. These conditions often result in severe motor impairments, limiting patients' mobility and independence. Brain-computer interfaces (BCIs) have emerged as promising tools for neurorehabilitation by translating neural signals directly into control commands for assistive devices. This thesis presents the development of a modular neurorehabilitation system that integrates BCI technology with real-time EEG signal processing, a Unity-based virtual front-end with its VR complement, and robotic arm control.

2 Proposed BCI System

The proposed neurorehabilitation system is a modular BCI platform designed to support motor recovery by linking real-time EEG analysis with interactive feedback and robotic actuation. EEG signals are acquired using an OpenBCI Cyton board, sampling at 1000 Hz. These signals are streamed to a host computer and processed through a pipeline that includes filtering, artefact removal, and feature extraction, such as Common Spatial Patterns (CSP) and band-power calculations. Two classification models are implemented: a convolutional neural network (CNN) trained directly on preprocessed EEG epochs, and a multilayer perceptron (MLP) trained on extracted features. A meta-classifier ensemble combines predictions from both models to improve accuracy. The Unity-based front-end and its VR complement provide visual feedback and engagement through a simulated rehabilitation environment. The front-ends communicate with a robotic arm using the FourMotors software, enabling physical interaction and proprioceptive reinforcement. To support training and evaluation, a custom dataset was created in a structured rehabilitation scenario involving cued motor imagery tasks. The overall system in Figure 1 forms a closed-loop control scheme, from neural intention through classification to visual and robotic feedback.

3 Results and Conclusion

Offline, both the CNN and the MLP implemented classifiers achieved strong subject-specific performance, with accuracies exceeding 75% on average. In real-time conditions, performance decreased, with the CNN reaching up to 56.9% accuracy and the MLP slightly higher at 57.3%, along with an F1-score of 0.72. For unseen subjects, all models showed a notable drop, with accuracies falling below 47%. The meta-classifier, combining both of them, did not

¹ student of the master's degree program Informatics and Its Specializations, field of study Medical Informatics, e-mail: tkment@students.zcu.cz

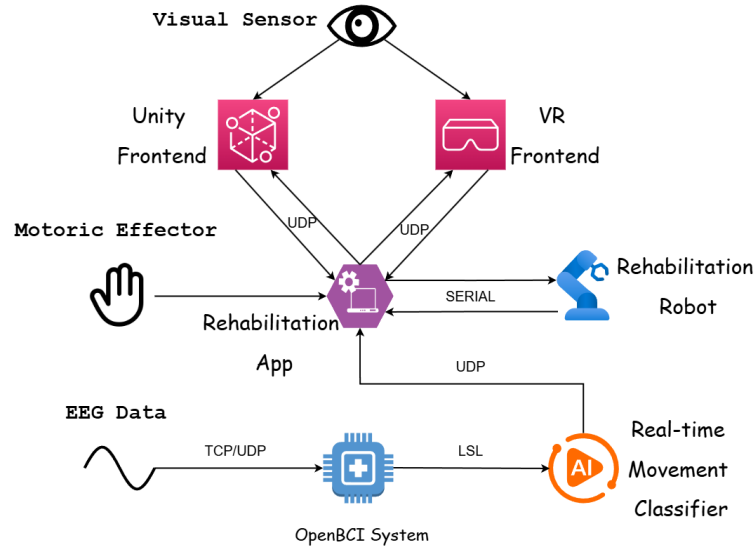


Figure 1: Diagram of proposed BCI system architecture

improve overall results and was biased toward predicting rest. The results, mentioned and depicted in Figure 2, showed promising accuracy for seen subjects, but also revealed significant generalization challenges when applied to unseen individuals, underscoring the difficulty of subject-independent EEG classification in real-world BCI scenarios. Despite these challenges, this thesis successfully delivered a modular and extensible platform for neurorehabilitation research. The insights gained during development, especially regarding EEG preprocessing, subject variability, and classifier design, provide a solid foundation for further work.

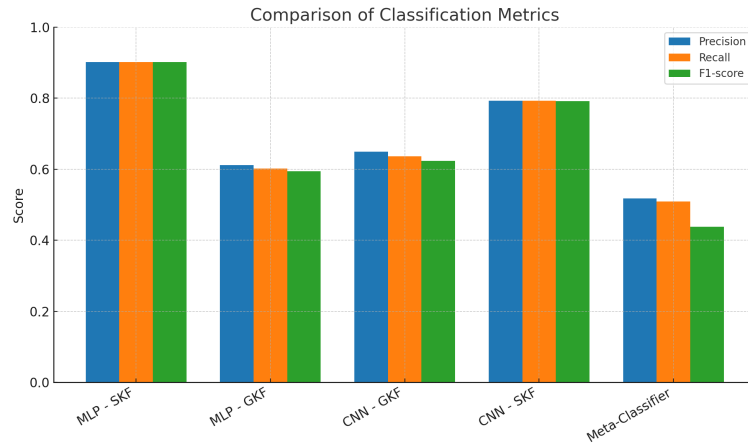
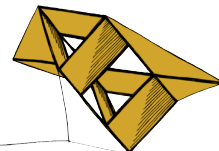


Figure 2: Precision, Recall, and F1-Score Across Model Setups

References

Feigin, V. L., Brainin, M., Norrving, B., Martins, S. O., Pandian, J., Lindsay, P., Gruper, M. F., & Rautalin, I. (2025) World Stroke Organization: Global Stroke Fact Sheet 2025. *International Journal of Stroke*, **20**(2), 132–144. <https://doi.org/10.1177/17474930241308142>.



Detekce a prevence driftu konfigurace síťových komponent v síti WEBnet

Aneta Koldovská¹

1 Úvod

V prostředí moderních počítačových sítí představuje správná a konzistentní konfigurace síťových zařízení klíčový předpoklad pro zajištění bezpečnosti, dostupnosti a správné funkčnosti celé infrastruktury. Jak zmínil Petryschuk (2022), v praxi často dochází k jevu známému jako *drift konfigurace* – tedy postupné odchylce aktuální konfigurace zařízení od jejího původně definovaného stavu. Tento drift může vznikat v důsledku manuálních zásahů či chyb při aktualizacích. Důsledkem může být zvýšené riziko výpadků, bezpečnostních incidentů nebo narušení síťových politik.

Tato práce se zaměřuje na problematiku detekce a prevence konfiguračního driftu v síťových komponentách v univerzitní síti WEBnet s cílem poskytnout systém, který lze využít ke včasnému odhalení změn. V rámci řešení je kladen důraz na automatizaci a integraci do běžných provozních procesů. Výsledky jsou demonstrovány na konkrétním laboratorním prostředí a vyhodnoceny z hlediska přesnosti detekce, nároků na správu a možností praktického nasazení.

2 Popis detekčního systému

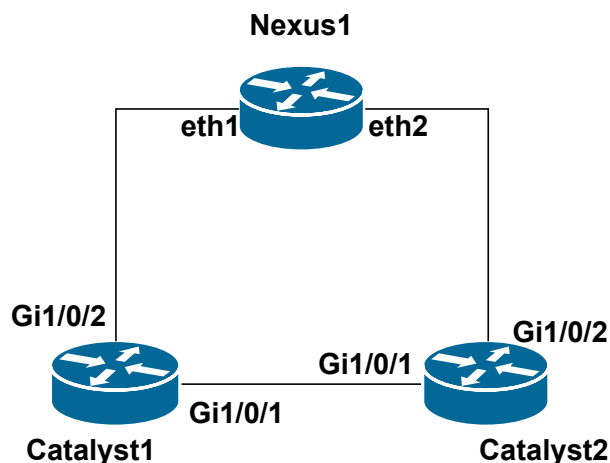
Navržený detekční systém je rozdělen do dvou hlavních částí, které společně zajišťují validaci výchozí konfigurace a následnou detekci případného driftu v aktuálně používaném nastavení síťových zařízení.

První část systému je integrována do CI/CD pipeline v prostředí GitLab a využívá nástroj Open Policy Agent (OPA), který umožňuje definovat a vyhodnocovat sadu pravidel nad výchozí konfigurací. Po nahrání nové konfigurace do příslušného GitLab repozitáře je automaticky spuštěna pipeline, která pomocí OPA nástroje ověří, zda konfigurace splňuje definovaná pravidla. V případě, že konfigurace pravidlům vyhovuje, je automaticky vytvořena nová verze a uložena jako nový výchozí referenční stav. Pokud pravidla nejsou splněna, je o tom neprodleně informován odpovědný síťový administrátor prostřednictvím e-mailového upozornění.

Druhá část systému slouží k detekci driftu aktuální konfigurace zařízení. V pravidelných intervalech je ze síťových prvků testovacího prostředí – vytvořeného pomocí nástroje *Containerlab* – automaticky stažena aktuální konfigurace. Topologie testovacího prostředí je zobrazena na obrázku 1.

Získané síťové nastavení je následně porovnáno s poslední platnou výchozí konfigurací, která prošla pravidlovou kontrolou. Pokud jsou v těchto konfiguracích nalezeny rozdíly, jsou považovány za konfigurační drift. Veškeré výsledky detekce jsou zaznamenávány do databáze a prezentovány prostřednictvím webového rozhraní dostupného na adrese <https://aether8.zcu.cz/>, kde lze vidět vývoj konfigurací jednotlivých prvků v čase.

¹ studentka navazujícího studijního programu Distribuované výpočetní systémy, specializace Distribuované systémy, e-mail: anetkold@students.zcu.cz



Obrázek 1: Schéma testovací topologie.

3 Závěr

Navržený detekční systém byl úspěšně otestován v kontrolovaném testovacím prostředí, kde byly do konfigurací síťových prvků úmyslně vkládány chyby a neautorizované změny. Systém dokázal tyto odchylky spolehlivě detekovat a správně je vyhodnotit jako konfigurační drift. Tyto testy potvrdily, že navržený nástroj je schopen spolehlivě detekovat změny v konfiguraci a poskytnout administrátorům užitečný přehled o vývoji nastavení síťové infrastruktury. Následně bylo přistoupeno k ověření funkčnosti na reálné infrastruktuře s fyzickými síťovými prvky.

Při testování v produkčním prostředí se však ukázalo, že některá zařízení kvůli svému stáří a omezené podpoře moderních protokolů neumožňují plnohodnotnou integraci detektoru. Z těchto důvodů nebylo možné systém v této infrastruktuře plně nasadit. Přesto bylo toto testování přínosné – odhalilo praktické limity nasazení takovýchto nástrojů v reálných podmínkách a zvýraznilo význam podpory otevřených standardů.

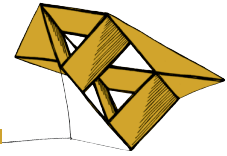
Zkušenosti z testování potvrdily, že pro efektivní detekci konfigurace driftu v produkčním prostředí je vhodné využívat moderní síťová zařízení, která podporují standardizované protokoly pro správu konfigurace (jako je například protokol NETCONF či RESTCONF).

Poděkování

Rád bych poděkovala vedoucímu této diplomové práce Ing. Maxmiliánovi Ottovi, Ph.D za jeho rady, ochotu a čas který mi věnoval při konzultacích. Také bych ráda poděkovala Ing. Martinovi Šimkovi, Ph.D za věcné připomínky a praktické rady při implementaci. Poděkování také patří Mgr. Petrovi Zimovi za zpřístupnění fyzických síťových prvků pro vyzkoušení provozu detektoru.

Literatura

Petryschuk, Steve. *What Is Configuration Drift and How to Avoid It* [online]. www.howtogeek.com, 2022 [cit. 2025/13/05]. Dostupné z: <https://www.auvik.com/franklyit/blog/network-configuration-drift/>



The Right Tracker for the Right Job: Maritime Case Study

Jan Krejčí¹

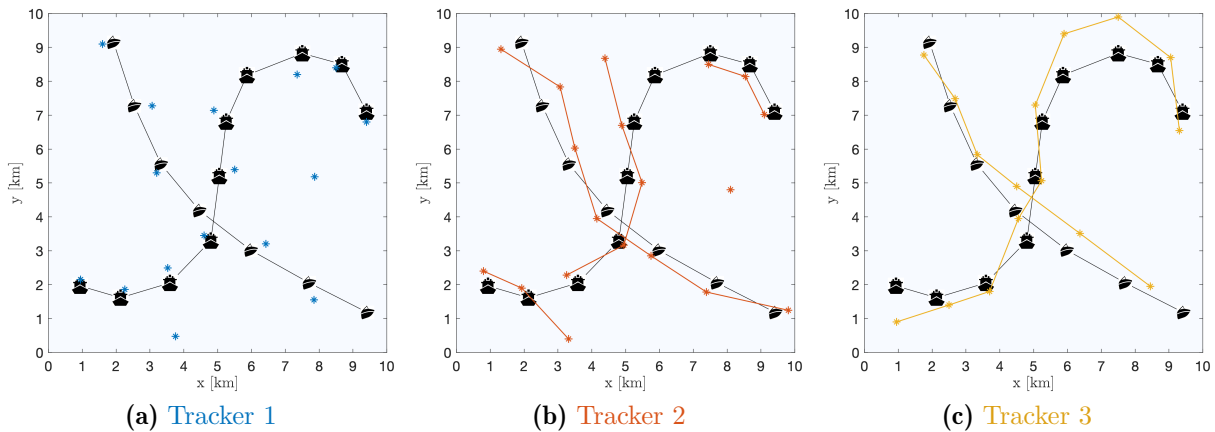


Figure 1: Three trackers are exemplified in a synthesized maritime scenario. Depending on the user's situation, each tracker can be best suited. *How to choose it properly?*

1 Introduction

Multi-object tracking refers to maintaining awareness of the positions/velocities of multiple moving objects. It is a key technology for naval or airborne applications, autonomous driving, or space situational awareness, among which many are safety-critical.

A large number of tracking algorithms, i.e., trackers, have been developed over the past decades, each having different complexity and performance. A few example trackers are illustrated in Figure 1 for a synthesized naval scenario, in which a speedboat (🚤) runs over a steamship vessel (🚢). A casual **user 1** may be a boat owner scanning for potential boats in his or her vicinity, for whom **Tracker 1** may suffice. Another **user 2** might be a watch tower operator requiring high-quality predictions, for whom **Tracker 2** may be more suitable. As the depicted scenario involves a hazardous situation, another **user 3** might need to analyze entire trajectories, hence **Tracker 3** may be the best choice.

The distinction between various trackers is seldom as clear as in the depicted scenario and practical scenarios may be rather complicated. To allow for *automatic* tracker selection, a user-adjustable performance evaluation metric is needed. Although safety-critical applications are often in question, such metrics relied on heuristics until recently.

2 Practical Performance Evaluation Metric

Recently developed *trajectory generalized optimal sub-pattern assignment* (TGOSPA) metric was shown to have many advantages over alternative metrics by García-Fernández et al. (2020). Moreover, the metric has several parameters whose practical selection was

¹student of the doct. deg. prog. Applied Sciences, field of study Cybernetics, e-mail: jkrejci@students.zcu.cz

discussed by Krejčí et al. (2024) for the case of visual tracking. This contribution extends the discussion by exploring the TGOSPA metric in the naval tracking example.

In particular, the TGOSPA metric parameters are the cut-off parameter c , power parameter p , and switching penalty γ . To achieve a simpler interpretation as depicted in Figure 2, Krejčí et al. (2024) suggested using the distance parameter a instead of p .

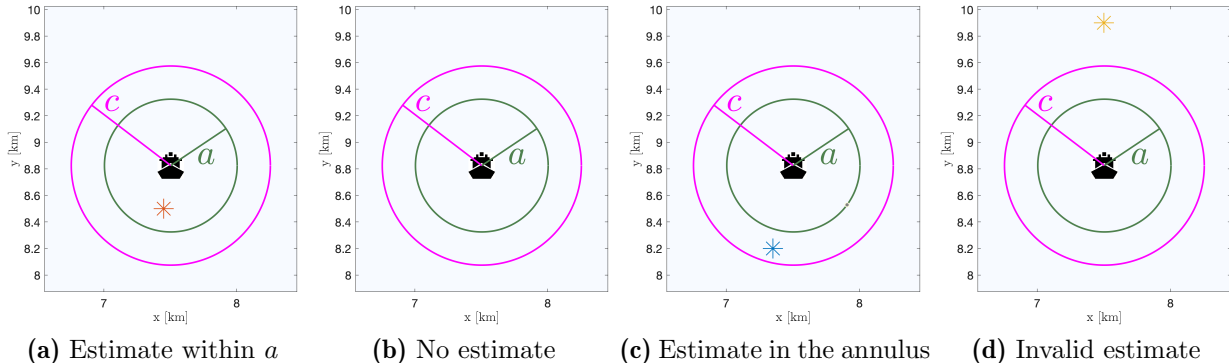


Figure 2: With the depicted parameters c and a , the TGOSPA metric would order the depicted estimates from left (best) to right (worst). The parameter γ then penalizes gaps and errors over entire trajectories which is harder to visualize, see Krejčí et al. (2024).

3 Results

The casual **user 1** is likely to prefer trackers having high localization precision over precise trajectory information, which suggests that such a user would select γ nearly zero. The watch-tower **user 2** may require some trajectory-related information rather than aiming for localization precision, suggesting that a slightly larger (small) γ may be preferable. To understand what happened over the entire scenario, the last **user 3** needs high-quality trajectory consistency over time, for which a large value of γ is best suited.

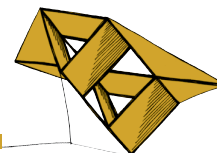
The values $\gamma=0.26$ km and $\gamma=0.75$ km for **user 2** and **user 3**, respectively, were selected based on guidelines from Krejčí et al. (2024). Results given in Table 1 are in accordance with the expert-based tracker suggestion discussed in the Introduction. Furthermore, the metric values can be given a sound meaning, see Krejčí et al. (2024).

Table 1: TGOSPA metric ($\downarrow$) with $c=0.75$ km, $a=0.5$ km and different parameter γ .

	γ zero	γ small	γ large
	user 1 aims for perfect localization, neglects trajectories	user 2 requires good localization with its history	user 3 aims for entire trajectories, sidelines localization
Tracker 1	2.41	2.76	3.63
Tracker 2	2.57	2.63	2.92
Tracker 3	2.77	2.85	2.88

References

- García-Fernández, Á., Rahmathullah, A. S., and Svensson, L. (2020) A Metric on the Space of Finite Sets of Trajectories for Evaluation of Multi-Target Tracking Algorithms. *IEEE Transactions on Signal Processing*, Vol. 68, pp. 3917–3928
- Krejčí, J., Kost, O., Straka, O., Xia, Y., Svensson, L., and García-Fernández, Á. (2024) *TGOSPA Metric Parameters Selection and Evaluation for Visual Multi-object Tracking*. Available from: <https://arxiv.org/abs/2412.08321> [Accessed 25th April 2025]



Engineering Thermostable Proteins Using Large Language Models

Lukáš Kuhajda¹, Tomáš Honzík², Jiří Vondrášek³, Daniel Georgiev⁴

1 Introduction

Proteins are functional units of organisms whose functioning is strongly dependent on temperature. Their sequences, made of 20 amino acids, can change behavior significantly with only a few mutations. These changes can shift the temperature range in which the protein is active.

Proteins support life of organisms from -20°C to $+122^{\circ}\text{C}$, while our proteins keep us alive in a range from $30\text{--}42^{\circ}\text{C}$. At low temperatures, enzyme activity drops due to reduced energy. At high temperatures, proteins can unfold and lose function permanently.

Proteins naturally adapt to temperature through mutations, but no clear rules guide how to design such changes. The relationship between protein structure, sequence, and function is too complex for manual or traditional computational methods. The global effort is huge to understand it and find new sources of proteins working in extreme conditions.

Proteins can be grouped based on their optimal working temperature: psychrophilic ($<20^{\circ}\text{C}$), mesophilic ($20\text{--}50^{\circ}\text{C}$), thermophilic ($50\text{--}80^{\circ}\text{C}$), and hyperthermophilic ($>80^{\circ}\text{C}$). Our project focuses on developing a system that can adjust proteins to shift their thermostability into a target category.

2 Methods

We built a new dataset of 3 million protein sequences. As a base model, we used Meta's ESM2 large language model (LLM) (Lin (2023)), pretrained on the task of sequence unmasking. We fine-tuned five separate models using our dataset to classify proteins into the four thermostability categories. The final classification is made by averaging the outputs of all five models.

To generate proteins with modified thermostability, we apply a genetic algorithm that introduces single-point mutations into a template protein. Each mutated variant is scored by the consensus of the five classifiers. This process continues until we find variants predicted to match the desired thermostability group.

3 Results

Our ensemble of five models was tested on public benchmark datasets and outperformed existing approaches in classifying proteins into non-thermophilic and thermophilic categories (Ahmed (2022), Charoenkwan (2022), Zhao (2023)).

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

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

³ Institute of Organic Chemistry and Biochemistry of the CAS, email:jiri.vondrasek@uochb.cas.cz

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

We used the system to generate multiple hyperthermophilic protein variants. All generated proteins retained the original 3D structure, verified using AlphaFold2 (Jumper (2021)). Simulations based on molecular physics confirmed that these variants have shifted thermostability as expected. Laboratory testing is currently ongoing at the Institute of Organic Chemistry and Biochemistry in Prague.

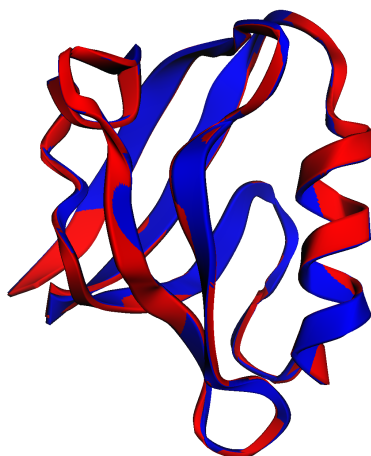


Figure 1: Comparison of a 3D structure: template protein (mesophilic) vs. generated protein (hyperthermophilic).

4 Conclusion

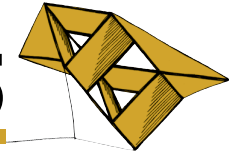
The fine-tuned models show strong classification performance and provide a reliable base for generating proteins with altered thermostability. When combined with a genetic algorithm, the system creates protein variants that maintain their structural integrity while adapting their thermal properties.

Acknowledgement

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

References

- Zeming Lin et al. ,Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*379,1123-1130(2023)
- Ahmed Z et al. iThermo: A Sequence-Based Model for Identifying Thermophilic Proteins Using a Multi-Feature Fusion Strategy. *Front Microbiol.* 2022 Feb 22;13:790063.
- Zhao J, Yan W, Yang Y. DeepTP: A Deep Learning Model for Thermophilic Protein Prediction. *Int J Mol Sci.* 2023 Jan 22;24(3):2217.
- Charoenkwan P, Schaduagratt N, Moni MA, Lio' P, Manavalan B, Shoombuatong W. SAP-PHIRE: A stacking-based ensemble learning framework for accurate prediction of thermophilic proteins. *Comput Biol Med.* 2022 Jul;146:105704.
- Jumper, John et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021, vol. 596, pp. 583–589



Analysis of White Matter Diffusion Properties in the Context of Selected Brain Tumors

Martin Kukrál¹

1 Introduction

Although brain tumors are rare, they cause mortality disproportionate to their incidence, with many histological types showing various demographic patterns and responsiveness to treatment (Leece et al., 2017). From the heterogenous group of brain tumors, malignant gliomas remain one of the most difficult types of tumors to treat, resulting in a multidisciplinary effort in designing new therapeutic paradigms to improve survival outcomes (Gunel et al., 2017).

This thesis explores the white matter diffusion properties around gliomas using the publicly available UCSF-PDGM dataset and an open-source computational pipeline.

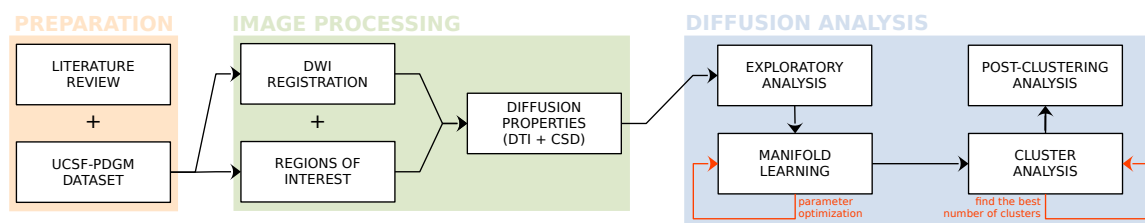


Figure 1: Thesis methodology overview

2 Image Processing

For each subject, the preprocessing stage consisted of nonlinear image registration of DWI data to the remaining MRI images, generation of peritumoral and periedematous ROIs using the operations of mathematical morphology, and the final computation of various diffusion properties via the DTI and CSD models. The resulting data consisted of aggregated diffusion properties in the peritumoral ($n = 421$) and periedematous ($n = 493$) ROIs.

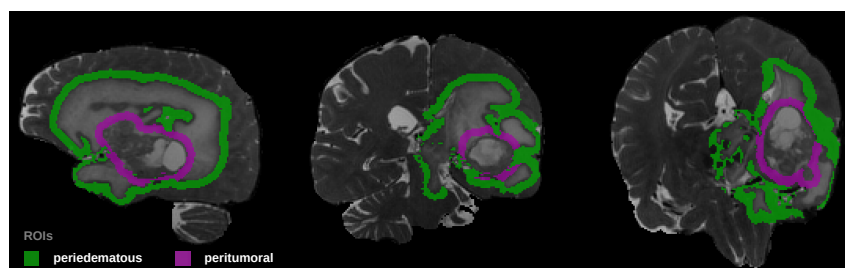


Figure 2: Peritumoral and periedematous ROIs generated around a high-grade glioblastoma

¹ student of the master's degree program Informatics and Its Specializations, field of study Medical Informatics, e-mail: kukrma@students.zcu.cz

3 Analysis of Diffusion Properties

The first step of the analysis was an exploratory data analysis comprised of visualizations and correlation analysis, followed by a nonlinear dimensionality reduction using UMAP, which had hyperparameters optimized by TPE, yielding more reliable 2D embeddings. These were then used in the cluster analysis via GMMs to uncover the structural patterns in the peritumoral and periedematous regions. The two peritumoral embedding clusters were consistently identified using multiple cluster quality indicators, but the periedematous embeddings did not cluster reliably. The post-clustering analysis of the peritumoral clusters revealed some associations between cluster labels and clinical markers.

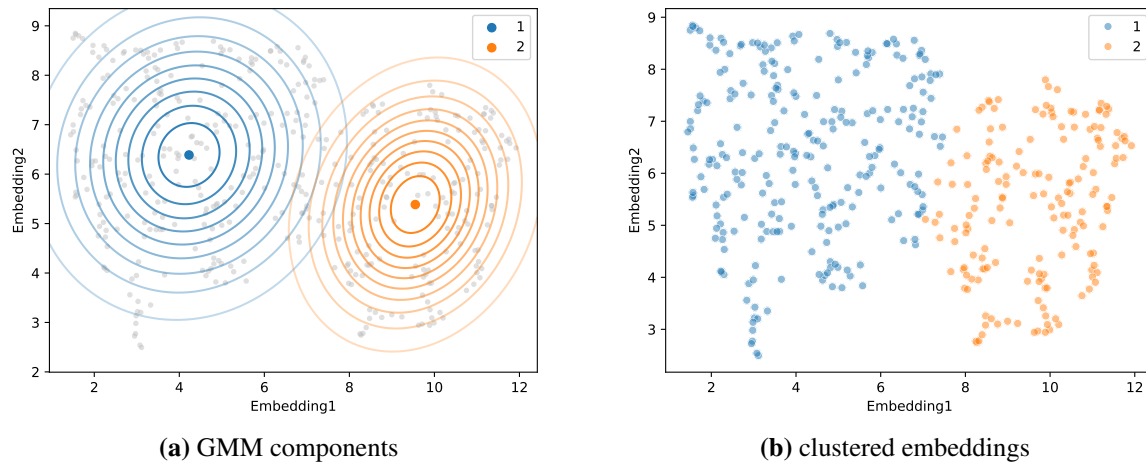


Figure 3: Cluster analysis of peritumoral UMAP embeddings

4 Results and Conclusion

The white matter diffusion properties in the peritumoral region are organized into two clusters corresponding to certain subtypes, which are weakly associated with patients' age, moderately to weakly with their sex, and very weakly with their overall survival. The periedematous properties do not cluster consistently.

Without additional clinical or biological markers, the clinical relevance of clusters remains unclear. These patterns may correspond to known but unavailable biomarkers or represent a previously unidentified glioma characteristic affecting surrounding white matter. Future works should focus on multimodal image analysis, explicitly incorporating the spatio-anatomical information, and including additional clinical and biological markers, all ideally in the context of low-grade gliomas due to their higher neurosurgical relevance.

References

- Leece, R., Xu, J., Ostrom, Q.T., Chen, Y., Kruchko, C., Barnholtz-Sloan, J.S. (2017) Global incidence of malignant brain and other central nervous system tumors by histology, 2003–2007. *Neuro-Oncology*. 19(11):1553–1564. ISSN 1523-5866. Available from: <https://doi.org/10.1093/neuonc/nox091>.
- Gunel, J.M., Piepmeier, J.M., Baehring, J.M. (2017) Preface. In: Gunel, J.M., Piepmeier, J.M., Baehring, J.M. (eds.). *Malignant Brain Tumors: State-of-the-Art Treatment*. Cham: Springer. pp. v–vi. ISBN 978-3-319-49863-8.

Zero Vibration Filter Design Tool

Martin Langmajer¹

1 Introduction

This paper presents a brief overview and description of an application developed as a comprehensive tool for designing a wide range of FIR input-shaping filters and their convolution. The application is available for the Matlab programming environment, as a stand-alone software for Windows operating system or as a virtual laboratory available at <https://virtual.pidlab.com/>.

The application *Zero Vibration Filter Design Tool* (ZVFDT) was created to provide a simple and intuitive interface for a synthesis of input shaping or generic FIR frequency filters. A graphical user interface (GUI) allows formulation of various design constraints, validation of important filter characteristics and combination of multiple filters using discrete convolution. The result of the design can then be exported to a .csv file in the form of filter impulse function. This allows its utilization in arbitrary real-time control or signal-processing platform.

2 Application options

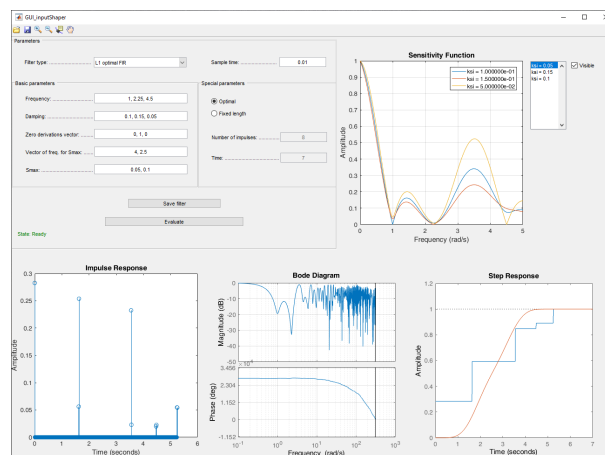


Figure 1: GUI of ZVFDT

The application allows synthesis of a wide spectrum of FIR filters. The main emphasis is placed on a class of Zero vibration (ZV) filters for the elimination of unwanted residual oscillations at specific frequencies. Both single and multi-mode filters are supported. Apart from that, generic frequency filters such as high-pass, low-pass, band-pass, or band-stop are supported as well. These filters can be used to selectively suppress a whole frequency band. The generic filters can be used separately, in a combination with a ZV filter or as a weighting function for specific ZV design methods allowing precise shaping of the filter frequency response.

¹ PhD student of Cybernetics, focused on control of robotic systems e-mail: gosh@students.zcu.cz

3 Usage example - Gantry crane

One of the particular problems concerning the motion control of gantry cranes is the flexibility of the rope which brings an inherent issue of hanging load oscillations which can be excited by the crane movement. Unwanted sway of the load during the process of manipulation may cause a collision with an obstacle or damage to the cargo. Precise motion control can be achieved by employing signal-shaping techniques.

Since there is usually no sensor measuring the sway angle, a convenient way to reduce the load oscillations is to use a setpoint shaping filter which processes the reference commands coming from a trajectory generator or from a human operator. Fig. 3 shows the experimental results obtained with the laboratory gantry crane model. The cart is driven by a three-phase AC induction motor controlled by a frequency inverter. Standard cascade PID position control is implemented using the feedback from the incremental encoder installed on the motor shaft. A four-pulse ZV4 input shaper which filters the position reference command was designed with the parameters chosen to achieve a suitable compromise between the filter duration and robustness. The plant parameters were obtained from the experimental identification. The load sway angle was not used as a feedback for the controller and it served only for the evaluation of the results. The cart was commanded to perform two point-to-point movements. It can be seen that the amount of excited oscillations of the hanging load was significantly reduce with the use of the shaping filter at the cost of slightly slower motion of the crane. A small level of residual vibration was still present due to the imperfect shape of the cart rail causing external disturbances which cannot be compensated in the feedforward manner.

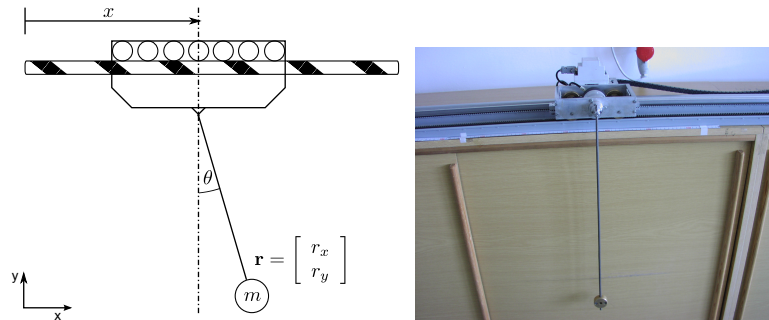


Figure 2: Gantry crane system - left kinematic setup, right experimental laboratory model

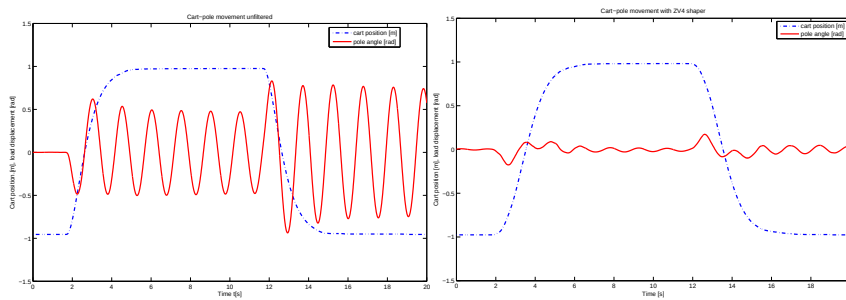
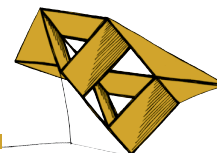


Figure 3: Motion control - left no signal filtering, right reference shaping using the ZV4 filter

Acknowledgement

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



Spatial Analysis of Image Descriptions for Detection of Cognitive Disorders

Tomáš Lebeda¹

1 Introduction

Neurodegenerative diseases pose a significant challenge to modern medicine, not only due to their progressive and irreversible nature but also because they often remain undiagnosed until symptoms become severe. Early detection is essential to maximize the effectiveness of available treatments and to slow the progression of cognitive decline. However, current diagnostic approaches are often slow, costly, and inaccessible, typically requiring in-person clinical assessments or expensive procedures like MRI.

Research by Bartoš et. al. (2024) has proposed an alternative approach that leverages cognitively demanding tasks - such as describing complex, detailed images in combination with other short-term memory tests as a means of early detection.

This paper explores the potential of automating the analysis of such image description tasks using machine learning techniques, with the goal of developing a fast, low-cost screening tool that can be used at home, potentially with the assistance of a family member, thereby reducing the need for in-person clinical visits. Specifically, the focus is on one specific sub-component of this approach: analyzing the sequence of objects as mentioned during image description.

2 Methodology

Participant is presented with an image (summer scenery on the riverbank with multiple people doing various activities) with the task to describe within one minute everything they see. The speech is recorded and later processed by automatic speech recognition (ASR) system to obtain text transcriptions, capturing word-level timestamps as well. These transcriptions are then processed using a custom-built semantic grammar parser (see Lebeda (2024)) and processing pipeline to extract structured semantic content from natural language.

Next, the mentioned objects are mapped to their spatial coordinates within the image, using a manually constructed reference description (structured JSON file, as described by Lebeda (2024)) as a ground truth. This allows us to track the sequence in which objects are described and their locations in the visual space.

From this sequence, we compute various spatial metrics, focusing primarily on the distances between successive object mentions. The underlying hypothesis is that individuals with cognitive impairment are more likely to produce disorganized, erratic descriptions, reflected in longer spatial jumps, while cognitively healthy individuals tend to describe the scene in a more systematic and spatially coherent manner.

¹ PhD student of Cybernetics, Artificial Intelligence; e-mail: lebedat@kky.zcu.cz

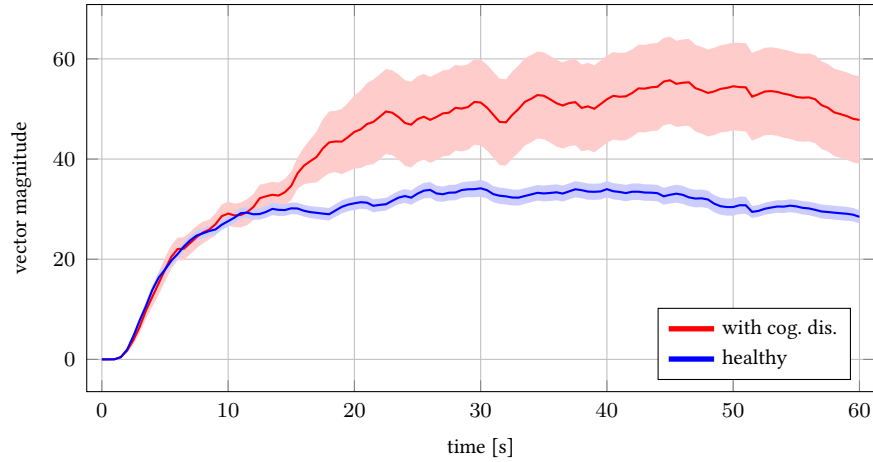


Figure 1: Mean magnitudes of vectors between successive object coordinates

3 Results and Conclusion

Figure 1 illustrates the magnitudes of vectors between successive object coordinates, across both participant groups. Initially, the average displacement is comparable between groups; however, as time progresses, individuals with cognitive impairment exhibit increasingly larger jumps, resulting in a greater mean vector magnitude.

The derived metrics were subsequently compiled into feature vectors of aggregated distance metrics over time segments and employed in classification tasks distinguishing the two groups, utilizing various machine learning classifiers, as presented in Table 1.

This approach leads to a promising results, though further empirical validation and refinement are needed.

Classifier	Precision	Recall	F1-score	Accuracy
Hist. Grad. Boosting	0.85	0.85	0.85	0.85
Logistic Regression	0.82	0.83	0.82	0.83
Random Forest	0.80	0.81	0.80	0.81
Nearest Neighbor (k=5)	0.87	0.85	0.81	0.85

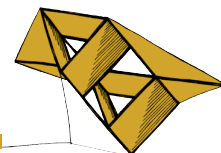
Table 1: Classification report

Acknowledgement

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

References

- Bartoš A, Zapletalová M, Zajícová L, Víta M, Šmídl L. *Popis obrázku k identifikaci kognitivní poruchy ve srovnání s certifikovanými testy Amnesia Light and Brief Assesment (ALBA) a Pojmenování obrázků a jejich vybavení (POBAV)*. *Cesk Slov Neurol N* 2024, 87, Supplementum 1, Sborník abstrakt, 37. český a slovenský neurologický sjezd, 2024, Ostrava
- LEBEDA, Tomáš. *Sémantická korespondence obrazu a řeči*. Online. Diplomová práce. Plzeň: Západočeská univerzita v Plzni, Fakulta aplikovaných věd. 2024. Dostupné z: <https://theses.cz/id/agk4a6/>.



Improving Sign Language Translation through Multimodal Language Alignment

Filip Majer¹

1 Introduction

Over 5% of the world's population experiences disabling hearing loss and many of them rely on sign languages as their primary means of communication. Despite the widespread use of sign languages, most modern communication technologies are designed primarily for spoken language, leaving deaf signers at a significant disadvantage. Recent advances in artificial intelligence offer new opportunities to address these communication barriers.

The main objective of this work was to design a new system for sign language translation. At its core is a novel video feature extraction model that combines both spatial and temporal information. In addition, the entire system supports pretraining through language alignment.

2 Background and Related Work

A central concept in this work is **language alignment**, which allows models to create similar representations for semantically related inputs across different modalities. This makes it possible for models like ChatGPT or Gemini to, for example, answer questions about images.

Another key component is the **DINO** framework, which enables self-supervised training of Vision Transformers and was introduced by Caron et al. (2021). This approach makes it possible to pretrain strong visual encoders without any labels and for instance, classify images it has never seen before.

3 System for Sign Language Alignment

The proposed system for sign language alignment is shown in Figure 1. The upper part processes video input and serves as the core model, **VideoDINO**. From each video frame, three regions of interest—face, left hand, and right hand—are extracted

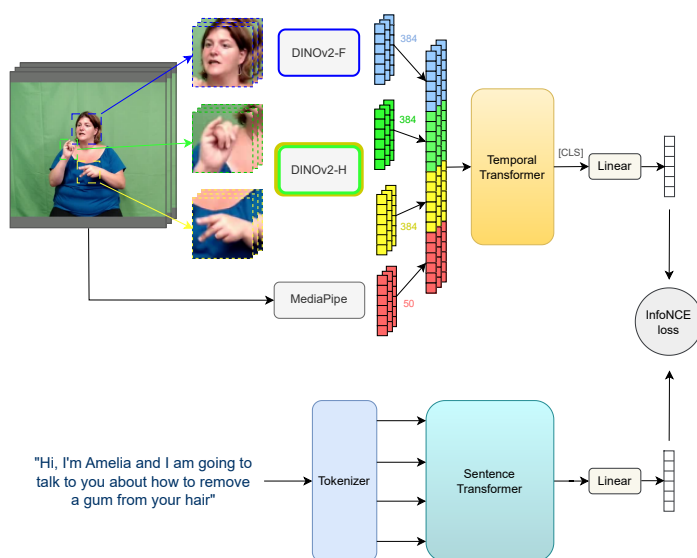


Figure 1: System for Sign Language Alignment.

¹ student navazujícího studijního programu Kybernetika a řídicí technika, specializace Umělá inteligence a automatizace, e-mail: fmajer@students.zcu.cz

and resized. These regions are encoded using separate DINOv2 models, which is an approach inspired by Gueuwou et al. (2024). In addition, 2D global pose keypoints are extracted using MediaPipe. The resulting vectors are concatenated and passed to a Temporal Transformer, which models temporal dependencies across the video sequence. The bottom part of the system contains a text encoder, which is used solely for aligning the video representations with textual transcription representations. Both video and text embeddings are projected into a shared space and trained with a contrastive InfoNCE loss.

4 Training

The system was trained on two datasets: How2Sign and YouTube-ASL. How2Sign offers high-quality, well-aligned transcriptions from studio recordings but has limited size and contains complex language. In contrast, YouTube-ASL is larger and features more day-to-day, conversational sign language, but with noisier and less consistent annotations.

5 Results

Figure 2 shows the results of the pre-training phase. Two histograms illustrate the cosine similarity distributions between video and text embeddings: one for matching pairs (positive samples) and one for mismatched pairs (negative samples). The separation between these distributions indicates successful sign language alignment.

For the translation task, the pretrained visual representations were used as input to a T5 model. As shown in Table 1, the proposed VideoDINO encoder significantly improved translation performance. Compared to the baseline without temporal modeling and alignment, a BLEU score increase of over 15% was achieved when using VideoDINO pre-trained on YouTube-ASL.

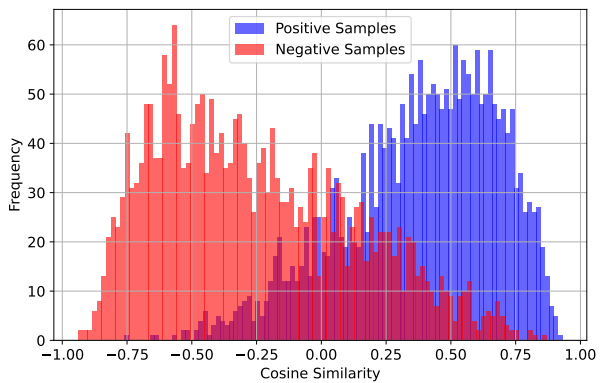


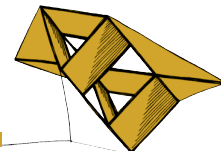
Figure 2: Cosine similarity histogram.

Visual Encoders	BLEU Score
Baseline (without temporal modeling and alignment)	0.91
VideoDINO - pretrained on How2Sign	0.60
VideoDINO - pretrained on YouTubeASL	1.05

Table 1: BLEU scores for sign language translation using different visual encoders.

References

- Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging Properties in Self-Supervised Vision Transformers. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9650–9660.
- Gueuwou, S., Du, X., Shakhnarovich, G., & Livescu, K. (2024). SignMusketeers: An Efficient Multi-Stream Approach for Sign Language Translation at Scale. *arXiv preprint arXiv:2406.06907*.



Two Pillars of Software Project Health: DevOps and DevEx

Elgun Majidov¹

1 Introduction

The success of modern software projects depends not only on technical excellence, but also on effective team collaboration and organizational alignment. Although many studies focus either on software engineering practices or human dynamics in isolation, this work presents a structured framework built on two pillars: DevOps and DevEx (Figure 1).

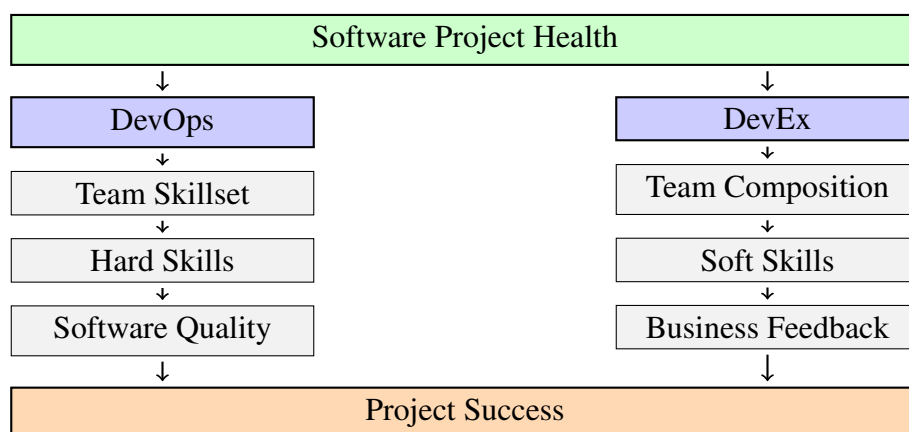


Figure 1: Framework of Software Project Health: DevOps and DevEx

The DevOps pillar includes hard technical skills, automation pipelines, and software quality assurance processes. It builds on research in continuous integration/deployment (CI/CD), secure delivery, and scalable systems. The DevEx pillar emphasizes soft skills, team composition, and stakeholder communication. It draws on studies on agile delivery challenges, business feedback loops, and cross-functional collaboration.

2 DevOps and Project Health

DevOps plays a key role in delivering reliable software efficiently. Forsgren et al. (2018) define four key metrics: deployment frequency, lead time for changes, time to restore service, and change failure rate, now widely used to assess delivery performance. Kula et al. (2022) highlight the role of team collaboration in meeting delivery timelines, while Zhu et al. (2020) stress the value of continuous feedback and monitoring. Based on these insights, potential indicators for DevOps health include:

- Deployment frequency and lead time,
- Change failure rate and recovery time,

¹ Student of the doctoral degree program Computer Science and Engineering, e-mail: emajidov@fav.zcu.cz

- Monitoring coverage and CI/CD automation levels,
- Handoff delays between teams.

Such data can be derived from CI/CD pipelines, incident logs, and source control systems.

3 DevEx and Project Health

DevEx focuses on the developer’s daily experience. Ernst et al. (2023) suggest minimizing cognitive load and improving feedback loops. Kalliamvakou et al. (2021) emphasize the inclusion of psychological and support factors in productivity models. Garbajosa et al. (2014) point out that communication issues in distributed teams can reduce DevEx and cohesion. Emerging DevEx indicators may include:

- Developer satisfaction and onboarding time,
- Interrupt frequency and feedback delay,
- Pull request turnaround and collaboration support.

These can be measured using internal surveys, issue tracking tools, and communication tools.

4 Conclusion

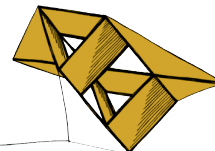
This work proposes a framework for evaluating software project health through two pillars: DevOps and DevEx. Moving beyond literature review, we identified practical indicators for each, grounded in real-world project data. Together, these metrics offer a basis for continuous, data-driven evaluation of both technical and experiential dimensions of software development.

Acknowledgement

The work was supported by the UWB grant SGS-2025-022 - New Data Processing Methods in Current Areas of Computer Science.

References

- Ernst, N.A., Bellamy, R.K.E. and Murphy-Hill, E., 2023. Developer experience: What actually drives productivity? *Communications of the ACM*, 66(2), pp.78–87. <https://doi.org/10.1145/3610285>
- Forsgren, N., Humble, J. and Kim, G., 2018. *Accelerate: The Science of Lean Software and DevOps*. IT Revolution Press.
- Kalliamvakou, E., Bird, C., Zimmermann, T. and Begel, A., 2021. Measuring developer productivity: Challenges and future directions. *Communications of the ACM*, 64(5), pp.72–79. <https://doi.org/10.1145/3449063>
- Garbajosa, J., Yague, A. and González, E. (2014) ‘Communication in agile global software development: An exploratory study’, in *Proceedings of the 15th International Conference on Product-Focused Software Process Improvement (PROFES 2014)*, Stuttgart, Germany, October 2014, pp. 408–417. doi:10.1007/978-3-662-45550-0-41.



Testing Platform for Periodic Control

Tomáš Myslivec¹

1 Introduction

Industrial processes face inherent disturbances from material variations, environmental factors, and equipment wear. At the same time, modern manufacturing depends on precision control for applications like robotics, servo systems, and power electronics, demanding robust control solutions that maintain both productivity and quality.

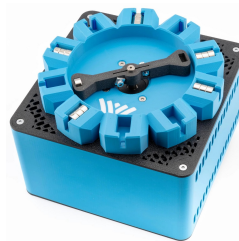
The presented Testing Platform for Periodic Control enables safe algorithm testing for systems with periodic disturbances. Its modular design allows easy reconfiguration for various disturbance patterns, while safety features like torque-limited motors and protective enclosures make it ideal for education.

2 Platform design

The main goal was to develop a cost-effective motion platform capable of periodic movements while fulfilling several key requirements. The platform needed to maintain reasonable purchase and maintenance costs while featuring a simple, modular design for easy configuration. Safety, robustness, and user-friendly operation were essential priorities in the design. Additionally, the system was intended to support a wide variety of laboratory exercises focused on compensating for periodic disturbances, making it versatile for different educational applications.



(a) Upper Ring



(b) The whole model

Figure 1: Periodic Motor Stand Model

The system integrates specialized software tools to support the development of efficient control algorithms, forming a complete periodic control solution. The entire construction is divided into two distinct parts: a box containing the motor and electronic components and an upper ring with slots for neodymium permanent magnets and a motor lever (see Fig. 1). This design allows for quick replacement of the upper ring with another one and thus future modification of the model. The motor lever has two sides that can be utilized, for instance, to

¹ student of the doctoral degree program Applied Sciences, field of study Cybernetics, e-mail: tmyslive@ntis.zcu.cz

double the number of peaks in disturbance signal within one period.

3 Experiment

To design a controller suppressing periodic disturbances within a given period, we will use the method of gradual controller improvement described in Brabec (2023). This approach is applicable to various types of controllers, but in this case, we specifically utilize PI and multiple PR controllers in parallel connection. Standard form of proportional resonant controller is denoted by

$$C_{PR}(s) = k_p + k_r \frac{2\omega_c s}{s^2 + 2\omega_c s + \omega_0^2}, \quad (1)$$

where ω_0 is the frequency of the harmonic to be suppressed, ω_c , is the design parameter selected by the designer.

4 Results

The final Fig. 2 presents the amplitude frequency response of the closed loop for each iteration and also output of the motor - velocity. The antiresonance peaks indicate the precise frequencies that are attenuated by incorporating PR controllers. The most substantial difference in velocity error is after adding the first PR controller, that suppress main harmonic frequency at 3.14 rad/s .

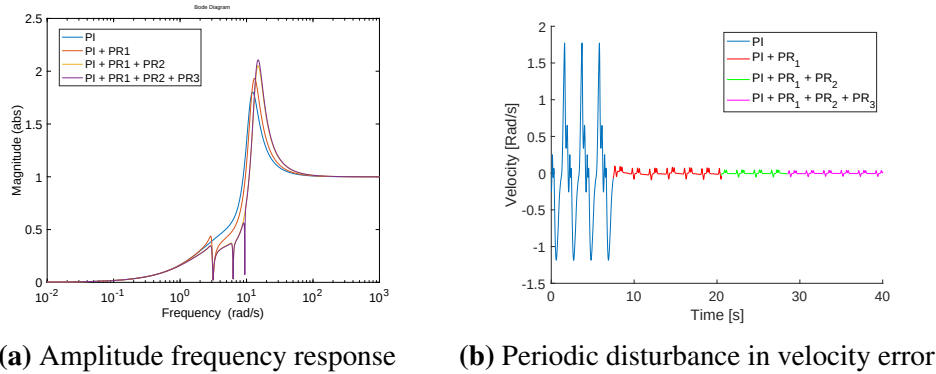


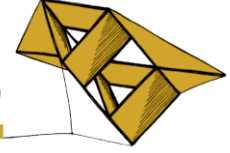
Figure 2: Final results

Acknowledgement

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

References

- Tomáš Myslivec and Miloš Schlegel (2025), "Education Platform for Periodic Control", 2025 25th International Conference on Process Control. (submitted)
- M. Brabec and M. Schlegel (2023), "Analytical Design of a Wide Class of Controllers with Two Tunable Parameters Based on H_∞ Specifications," 2023 24th International Conference on Process Control (PC), Strbske Pleso, Slovakia, 2023, pp. 221-226, doi: 10.1109/PC58330.2023.10217459.



SpiTranNet: A Spiking-Transformer network for Motor Imagery classification using multi-channel EEG signals

Duc Thien Pham¹

1 Introduction

Brain-Computer Interfaces (BCI) translate brain signals into commands, enabling control without muscle movement, which is crucial for motor-impaired users. Electroencephalography (EEG) is commonly used in BCIs because it is easy to use, despite low spatial resolution. Motor Imagery (MI) changes brain signals (ERD/ERS), helping detect a person's movement intentions. In this study, we introduce SpiTranNet, a Spiking-Transformer model for MI classification using EEG signal. It handles binary and multiclass tasks, showing strong performance for BCI-based rehabilitation.

2 Method

We have developed the SpiTranNet model to classify MI binary and multiclass classification. The dataset was collected at the University of West Bohemia in Pilsen. The architecture of the proposed method is shown in Figure 1.

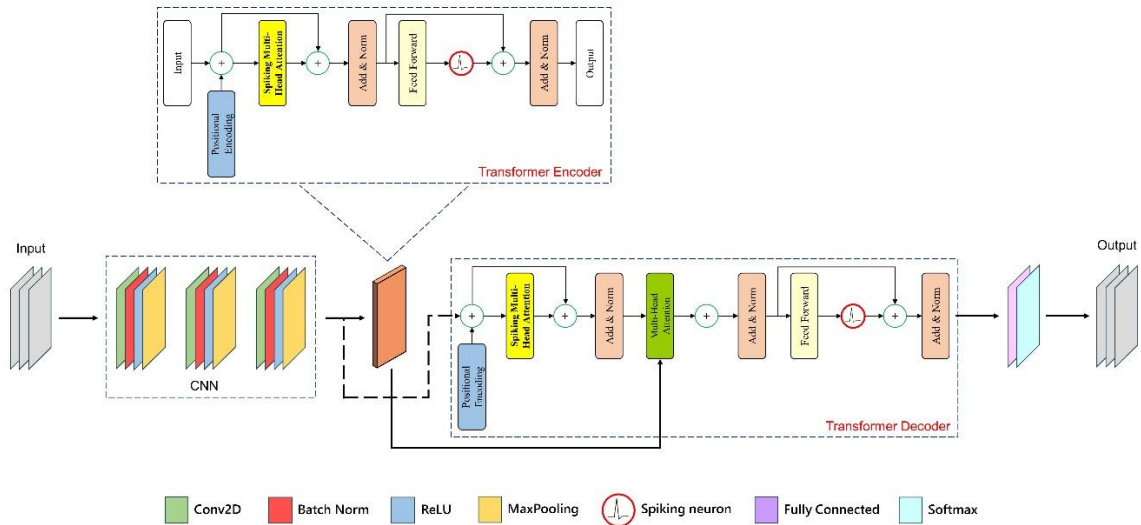


Figure 1: The architecture of the SpiTranNet model.

The Spiking neuron $S(t)$ can be approximated as:

$$S(t) \approx \sigma(V(t)) = \frac{1}{1 + \exp(-\text{temp} \cdot (V(t) - \text{threshold}))} \quad (1)$$

Then, we propose the Spiking Multi-Head Attention (SMHA) mechanism, which enhances traditional self-attention by incorporating spiking activations, enabling attention outputs to exhibit discrete event-driven processing similar to biological neurons.

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

$$SMHA = S(Attention(Q, K, V)) \quad (2)$$

3 Results

As shown in Tables 1 and 2, the SpiTranNet model outperforms state-of-the-art methods in both binary and multiclass MI classification tasks.

Methods	Accuracy	Precision	Recall	F1-score	AUC	k
SVM	57.28±0.47	57.28±0.47	57.26±0.47	57.24±0.47	-	-
LDA	50.51±0.99	50.53±0.99	50.52±0.99	50.49±0.98	-	-
MLP	51.12±1.49	51.16±1.55	51.13±1.52	50.55±1.90	-	-
CNN	76.00±0.80	76.73±0.75	76.05±0.79	75.86±0.90	-	-
EEG-ITNet	75.45±1.43	76.43±0.96	75.50±1.40	75.23±1.58	0.755±0.01	-
CNN-Transformer-LSTM	78.51±1.21	78.58±1.19	78.52±1.21	78.50±1.21	0.785±0.01	-
CNN-LSTM	79.06±1.47	79.13±1.41	79.07±1.47	79.05±1.48	0.791±0.01	-
CNN-Transformer	77.93±0.68	77.96±0.69	77.93±0.68	77.92±0.68	0.779±0.01	-
SpiTranNet	79.09±0.62	79.17±0.63	79.12±0.62	79.10±0.62	0.792±0.01	0.581

Table 1: Performance comparison with existing methods for binary classification

Methods	Accuracy	Precision	Recall	F1-score	AUC	k
SVM	40.35±0.82	40.43±0.82	39.82±0.79	39.48±0.80	-	-
LDA	35.26±1.46	35.34±1.45	35.32±1.47	35.20±1.42	-	-
MLP	32.94±1.13	33.03±1.21	33.00±1.21	32.49±1.03	-	-
CNN	58.57±1.45	60.32±1.48	58.43±1.48	57.99±1.72	-	-
CNN-Transformer-LSTM	61.95±1.99	62.73±2.11	61.84±2.07	61.80±2.03	0.715±0.02	-
SpiTranNet	62.11±1.04	62.75±1.12	61.96±1.05	61.99±1.05	0.716±0.01	0.427

Table 2: Performance comparison with existing methods for multiclass classification

4 Conclusion

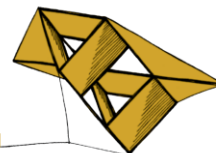
This study applies the SpiTranNet model to MI EEG dataset, achieving top accuracy of 79.09±0.62% (binary) and 62.11±1.04% (multiclass). By combining spiking neural networks and Transformers, the model effectively captures EEG patterns, showing strong potential for BCI applications and reducing human effort.

Acknowledgement

This work was supported by the University specific research project SGS-2025-022 New Data Processing Methods in Current Areas of Computer Science (project SGS-2025-022).

References

- Mouček, R., Koderá, J., Mautner, P., and Prucha, J. (2024). Augmentation of Motor Imagery Data for Brain-Controlled Robot-Assisted Rehabilitation. *Proceedings, BIOSTEC*, 2, pp. 812-819.
- Titkanlou, M. K., Monjezi, E., and Mouček, R. (2024). A Type of EEG-ITNet for Motor Imagery EEG Signal Classification. *Proceedings, BIOSTEC*, 2, pp. 257-262.
- Pham, D. T. and Mouček, R. (2024). Automatic Motor Imagery Classification by CNN-Transformer-LSTM Using Multi-Channel EEG. *Proceedings, ECAI 2024*, pp. 4555-4562.
- Titkanlou, M. K., Pham, D. T., and Mouček, R. (2025). Classification of EEG Signal Using Deep Learning Architectures Based Motor-Imagery for an Upper-Limb Rehabilitation Exoskeleton. *SN Computer Science*, 6(3), 193.



Segmentace cév prasete domácího ve snímcích z výpočetní tomografie

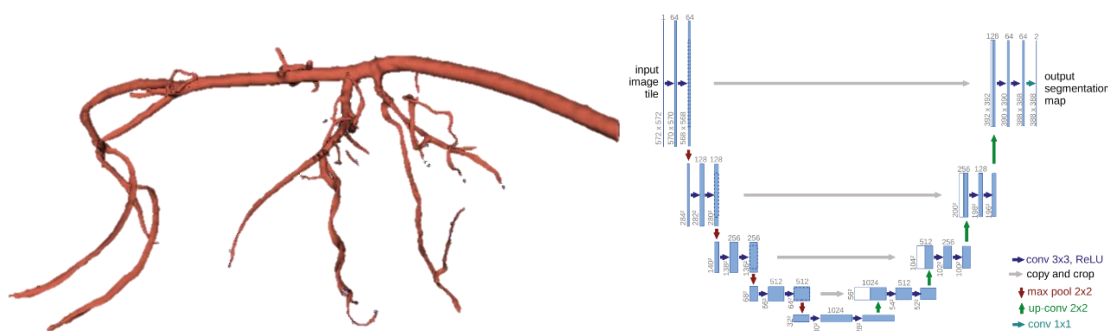
Jana Romová¹

1 Úvod

Cílem práce bylo navrhnout, otestovat a implementovat metodu pro segmentaci cév v dutině břišní prasete domácího, zejména v okolí jeho jater. Vyvinutá metoda by měla sloužit vědeckému týmu Univerzity Karlovy, který se zabývá resekčními zákroky na zvířecích játrech, jejich regeneračními schopnostmi a potenciálem pro budoucí použití při transplantacích u lidských pacientů s nenávratným poškozením jater.

2 Východiska

Pro zadanou úlohu dosud neexistují vhodné anatomické sady dat ani výzkum vhodných metod pro segmentaci zvířecích cév. Prvním krokem proto bylo vytvořit s pomocí pracovníků Lékařské fakulty vlastní dataset cévních stromů prasat domácích Pilsen Pigs (Obr.1), na kterém by bylo možno vyvíjenou metodu vyhodnotit. Protože anatomie prasečích jater je velmi podobná té lidské, byla jako vhodná architektura hledané metody zvolena síť U-Net (Obr. 1), která se pro podobné aplikace používá na lidských CT snímcích např. TotalSegmetnator.



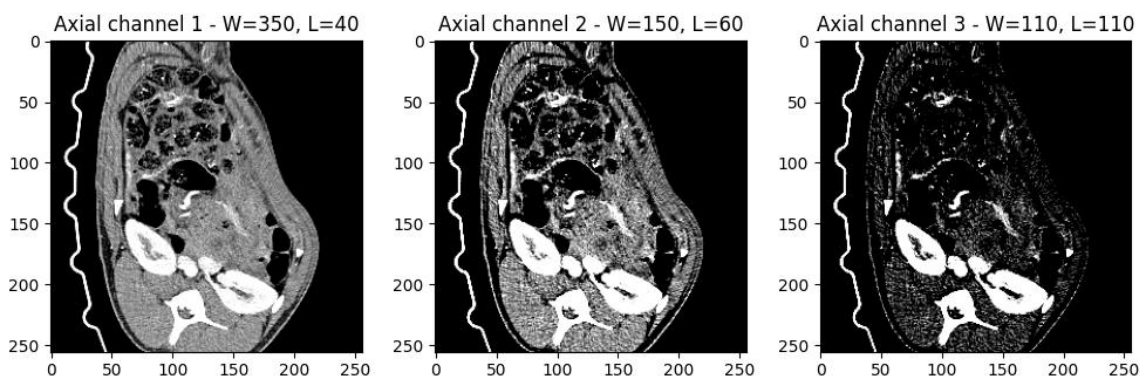
Obrázek: 1: Vlevo: segmetnovaný cévní strom jater prasete domácího z nového datasetu Pilsen Pigs. Vpravo: Klasická architektura U-Net vhodná pro segmetnace v medicínských snímcích MRI (magnetická rezonance) a CT (výpočetní tomografie).

Během vývoje metody byla analyzována 2D i 3D architektura sítě U-Net a její parametry. Jednotlivé modely byly testovány na veřejných datasetech Deepvesselnet (malé cévy) a datasetu 3D-Ircad (lidská játra a velké cévy). Analýza vedla k těmto závěrům:

- 2D verzi architektury U-Net není možno použít, ani po úpravě parametrů, protože dochází ke ztrátě 3D informace a malé cévy jsou chybně považovány za šum v obraze.
- Parametry 3D architektury je třeba výrazně pozměnit:

¹ studentka navazujícího studijního programu Informatika a její specializace, obor Medicínská informatika, e-mail: romova9@students.zcu.cz

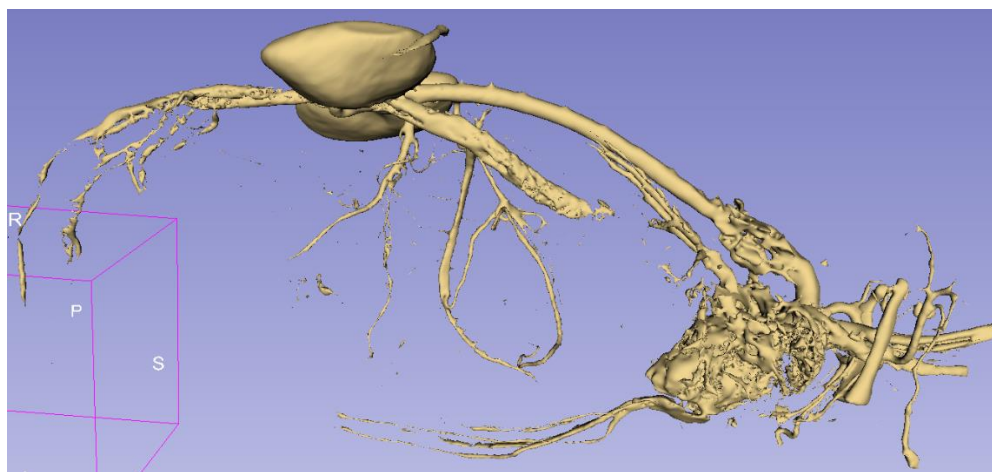
- Zvýšit hloubku a výšku konvolučních vrstev, aby byl model schopen zachytit jemné cévy.
- Použít váhové verze ztrátových funkcí pro vyrovnání zastoupení cév v celkovém objemu dat.
- Na vstupu modelu je nutno použít snímky ve třech různých variantách předzpracování, aby byly zachyceny různé anatomicky významné struktury břišní dutiny z původního snímku CT (Obr. 2).



Obrázek 2: Ukázka vstupů segmentačního modelu. Na vstupu jsou 3D matice s oříznutými intenzitami jasu. Zleva oříznutí zobrazuje podrobněji: orgány, svaly a prokrvené tkáně, kosti.

3 Výsledky

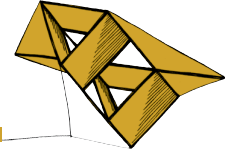
Pozměněná architektura byla nakonec otestována na nově vzniklém datasetu Pilsen Pigs, kde během validace dosahovala metrika IoU (Intersection over Union) hodnoty 0,97. Výsledná metoda byla implementována jako rozšíření aplikace 3DSlicer, která se běžně používá pro medicínské zobrazování.



Obrázek 3: Ukázka segmetnace novou metodou v aplikaci 3D Slicer.

Literatura

- Ronneberger, O., Fischer, P., a Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation.
- Soler, L., Hostettler, A., Agnus, V., Charnoz, A., Fasquel, J., Moreau, J., Osswald, A., Bouhadjar, M., a Marescaux, J. (2010) 3D image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database. *IRCAD*.



Computer Vision Applications in Video Recordings for Traffic Signal Detection and Classification on Czech Railways

Daniel Schnurpfeil¹

1 Introduction

Railway signaling systems serve as the safety control mechanism for locomotive drivers. That is an argument to introduce automated recognition of railway signals. The thesis focuses on automated visual detection and classification of railway signals within the Czech railway system using a combination of traditional computer vision techniques and modern deep learning architectures.

2 Czech Railways

Railway signaling systems contain a lot of information on speed restrictions and operational permissions. The Czech railways have a specific signaling system consisting of various signal types, such as signals with single or multiple lights, each with specific meanings and operational implications (Figure 1).

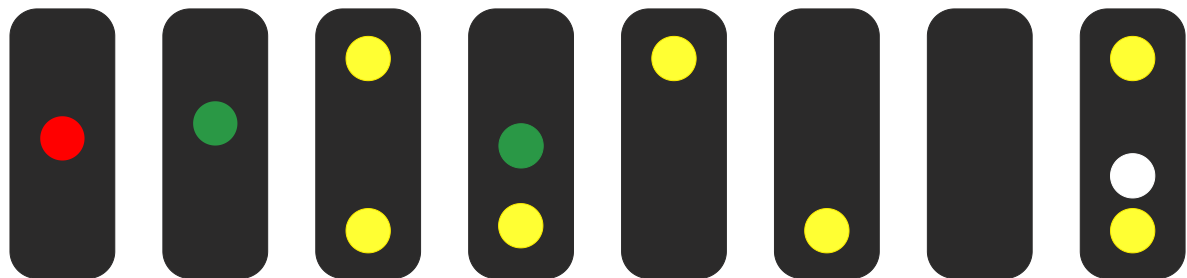


Figure 1: Stop Signal on the Left; Green Light; Limit 40 km/h and warning; Limit 60 km/h and go; warning; (also) warning; lights off; Speed 40 km/h and repeating the signal Warning on the right

3 Results

Experimentation was performed with various state-of-the-art object detection architectures: YOLOv10, YOLOv5, YOLOv8, RT-DETR-l and RT-DETR-x. The hyperparameter optimization strategy consisted of different freezing levels (0 and 3), training epochs (15, 30, 40, 60, 80, 100, 150 and 200), learning rate (0.0001, 0.001, 0.01) and confidence thresholds (0.3, 0.5, and 0.7) to determine the optimal configuration for the detection performance of the railway signal.

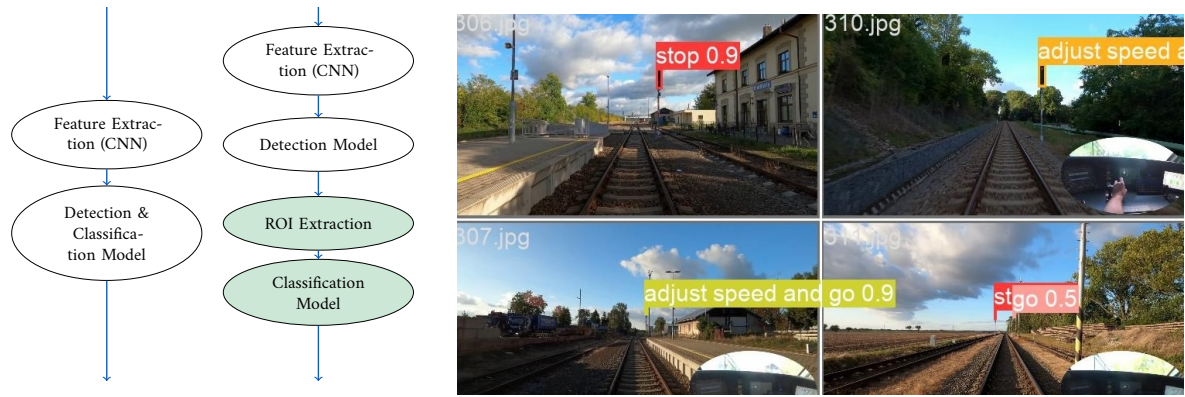
¹ student of the master degree program Software Engineering, field of study Computer Science and Engineering, specialization Natural Language Processing, e-mail: dschnurp@students.zcu.cz

Table 1: Best Grid Search Results

Model	Epochs	Lr	mAP50	mAP50-95	Precision	Recall	F1
YOLOv5mu	100	0.001	0.8277	0.7112	0.9459	0.6904	0.7983
YOLOv5nu	100	0.001	<u>0.8560</u>	0.7243	0.9113	<u>0.7934</u>	0.8486
YOLOv10n	100	0.001	0.8626	0.6891	0.8667	0.7586	0.8090
YOLOv8n	80	0.0001	0.8076	0.6866	0.8795	0.7119	0.7871
RT-DETR-L	60	0.001	0.8230	0.6793	0.8734	0.8049	0.8378
RT-DETR-L	100	0.001	0.7999	0.6424	0.8987	0.7911	<u>0.8414</u>

4 Conclusion

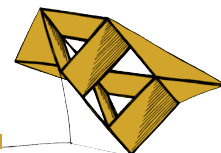
A multistage approach integrated YOLO/RT-DETR models for signal detection, followed by CNNs for state classification. The research introduced a semiautomatic annotation pipeline for train cabin videos that significantly enhances dataset creation efficiency (final dataset with over 2000 samples). Experimental results demonstrate that railway signals can be detected with a high mean Average Precision (mAP50-95 0.72 valid) (Table 1) and F1 score 0.9 (test). The developed system can distinguish between critical signal states including stop, go, warning, and various speed adjustment indicators.

**Figure 2:** Single-stage vs Two-stage Model; Detection Preview

In future research directions, the detection of flashing lights and repeating signals with white light (Fig. 1 right) could be included, which goes beyond the scope of this thesis and requires a more complex analysis and additional data collection. The integration of the proposed system into onboard cameras should need accelerated inference speeds to adapt to the high velocity of moving trains. Detection performance was poor, when signals were placed against complex backgrounds where the color contrast was lower and by challenging illumination conditions, also noted by Staino et al., 2022.

References

Staino, A., Suwalka, A., Mitra, P., & Basu, B. (2022). Real-time detection and recognition of railway traffic signals using deep learning. *Journal of Big Data Analytics in Transportation*, 4, 57–71. <https://doi.org/10.1007/s42421-022-00054-7>



Modelování a identifikace portálových jeřábů

Daniel Sukovatý¹

1 Úvod

V dnešní době jeřáby představují klíčový prvek především ve stavebnictví a průmyslu. Jejich použití velmi přispívá k úspěšnému, bezpečnému a rychlému dokončení úkolů. Dnešní svět klade velký důraz na automatizaci a to například z důvodu zvýšení rychlosti dokončení dané části procesu nebo redukce chybovosti způsobené lidským faktorem. Právě lidský faktor hraje velkou roli v manipulaci s jeřáby.

Automatické tlumení nežádoucích kmitů by velmi zjednodušilo a zrychlilo ovládání. Návrh řízení vyžaduje znalost matematického modelu a každá změna parametrů jeřábu změni i optimální nastavení regulátoru. Parametry lze jednorázově změřit v případě lana a háku, ale pro zátěž by to byl zdoluhavý a finančně náročný proces. Proto je zapotřebí metody, která dokáže odhadovat tyto parametry během provozu jeřábu. Vzhledem k časté výměně zátěže je důležité, aby metody byly časově nenáročné.

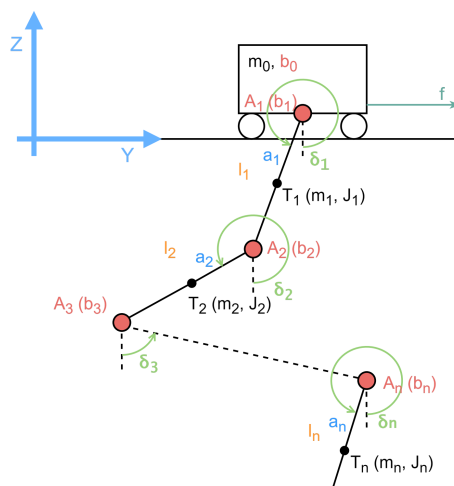
Tato práce se zabývá metodami pro identifikaci neznámých parametrů zátěže, které jsou následně použity pro návrh optimálního regulátoru.

2 Odvození modelu

Jako model je uvažováno trojitě kyvadlo zavěšené na posuvném vozíku. Každé rameno (resp. kloub) disponuje vlastními fyzikální parametry, kterými jsou (viz. obr. 1):

- a_i - poloha těžiště,
- l_i - délka ramene,
- m_i - hmotnost,
- J_i - moment setrvačnosti,
- b_i - koeficient tlumení.

Odvození je založeno na definování pohybových rovnic pomocí Lagrangeovy metody¹. Následné pohybové rovnice je zapotřebí linearizovat kolem rovnovážného bodu, kterým je v tomto případě natočení kloubů v dolní úvratí. Výsledkem linearizace je model ve stavovém popisu.

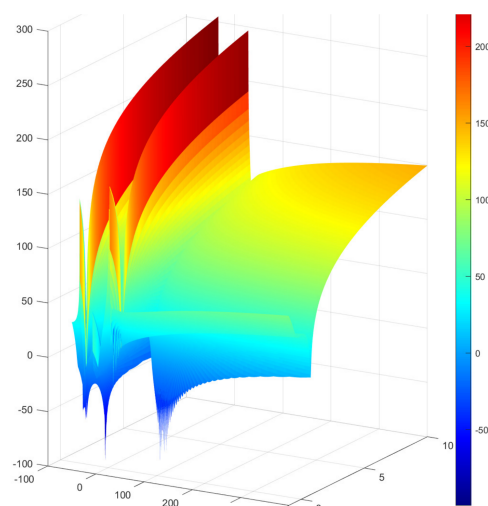


Obrázek 1: Obecné schéma kyvadla s n -rameny

¹ student navazujícího studijního programu Kybernetika a řídicí technika specializace Automatické řízení a robotika a výzkumný pracovník centra NTIS, e-mail: dansuk24@students.zcu.cz

3 Identifikační metoda

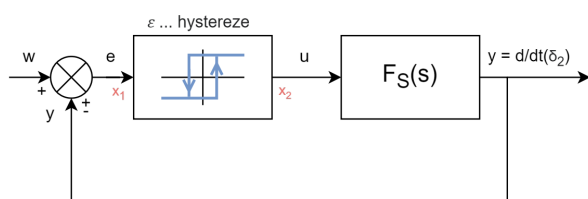
V této části je předpokládána znalost prvních dvou rezonančních frekvencí systému. Tato metoda je založená na porovnávání charakteristických polynomů matice dynamiky a její ekvivalentní Jordanovy formy. Porovnáním vznikne soustava šesti rovnic, ze kterých následnými úpravami a dosazeními se lze dostat na soustavu o dvou rovnicích. Z těchto rovnic je možné sestavit účelovou funkci (viz. obr. 2), jejíž argumentem minima jsou předpokládané hodnoty fyzikálních parametrů. K nalezení minima byla použita negradientní optimalizační metoda *Pattern Search*².



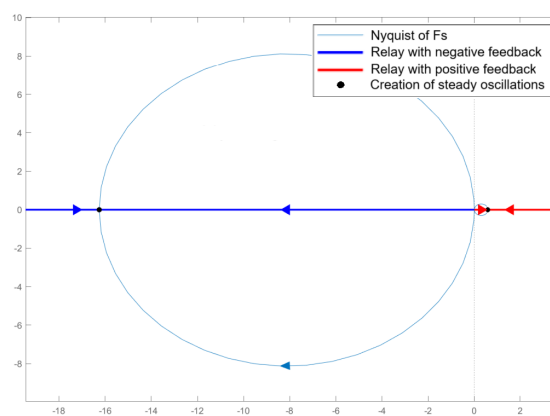
Obrázek 2: Účelová funkce

4 Odhad rezonančních frekvencí

Tato část se věnuje odhadu prvních dvou rezonančních frekvencí systému, které jsou následně využity k výpočtu chybějících parametrů. Podle metody harmonické linearizace³ platí, že zapojením relé do zpětné vazby budou vybudeny kmity o frekvenci, která odpovídá právě jedné z rezonancí (viz. obr. 3). Volbou znaménka zpětné vazby pak lze zvolit, zda bude odhadována první nebo druhá rezonance (viz. obr. 4).



Obrázek 3: Schéma uzavřené smyčky



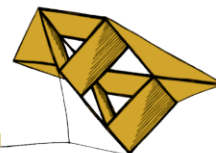
Obrázek 4: Nyquistův diagram

5 Závěr

Cílem práce bylo navrhnout metody, které na základě známé struktury modelu a fyzikálních parametrů lana a háku umožní odhadnout zbývající parametry zátěže. V rámci simulací vedly navržené postupy k výsledkům s dostatečnou přesností. Při testování na reálném systému se však vyskytly významné odchylky v odhadu rezonancí, zejména u druhé rezonance, což vedlo k nepřesným hodnotám odhadovaných fyzikálních parametrů. Dalším požadavkem byla nízká časová náročnost, která byla z pohledu odhadu rezonancí splněna, zatím co metoda identifikace parametrů trvala řádově několik minut. Doporučením pro další výzkum je analýza nepřesností mezi modelem a reálným systémem nebo využití alternativní metody identifikace, například online odhadu pomocí Kalmanova filtru.

Literatura

1. Königsmarková, J., Schlegel, M. (2017) *Identification of n-link inverted pendulum on a cart*. NTIS, Západočeská univerzita v Plzni.
2. Neri, F., Rostami, S. (2021). *Generalised pattern search based on covariance matrix diagonalisation*. SN Computer Science, 2, Article 171. <https://doi.org/10.1007/s42979-021-00513-y>
3. Blaha, L., Kubíček, K. (2023). *Nelineární systémy*. Západočeská univerzita v Plzni.
4. Sukovatý D. (2024). Modelování a identifikace portálových jeřábů (Bakalářská práce, Západočeská univerzita v Plzni). <https://dSPACE.zcu.cz/items/22d5664e-e6fc-4e3f-b92b-3cd3ad231e89>



Klasifikace akustických událostí pro diagnostiku výrobní linky

Jan Šimek¹

1 Úvod

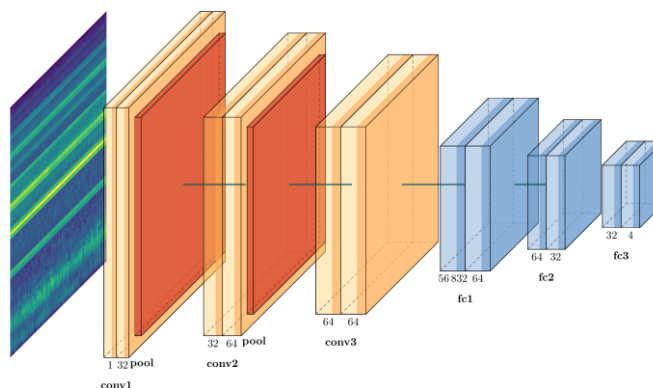
Zkušeni zaměstnanci výrobní linky závodu poskytují cennou zpětnou vazbu o stavu jejího strojního vybavení. Jsou mimo jiné schopni pouze na základě sluchu rozpoznat počínající závady strojů a včas je hlásit technikům. V době postupné automatizace a snižování počtu osob v provozu však tento lidský dohled slábne, což vyvolává obavy ze snížení obecného povědomí o aktuálním dění na lince. Tato práce si klade za cíl ověřit, zda lze navrhnout neurální model, jenž na základě zvukových signálů rozpozná klíčové stavy jednotlivých strojů provozu a správně je klasifikuje i v abnormálně hlasitém prostředí. Výsledkem bude první krok k vývoji detektoru nežádoucích akustických událostí, který doplní lidský dohled v modernizovaném provozu.

2 Charakteristika dat

Pro účely práce byla vytipována dvě zařízení vhodná k nahrávání: čerpadlo a zásobník na slad. Čerpadlo má 3 čerpací režimy lišící se jeho výkonem (60, 90 a 70 %), které v proceduře trvají 3, 7 a 5 minut a vždy se bezprostředně následují ve stejném pořadí, přičemž celá procedura se opakuje v intervalu okolo dvou hodin. U zásobníku na slad detekujeme pouze samotné vysypání sladu do zásobníku, které trvá zhruba 3 sekundy a během procedury plnění se opakuje každých 10 sekund. Naplnění zásobníku takto trvá přibližně hodinu a půl.

3 Architektura modelu

Architektura použité konvoluční sítě (Nandi, 2021) je znázorněna na obrázku 1. Do sítě vstupuje spektrogram, následují 3 konvoluční vrstvy (32, 64 a 64 jader velikosti 3x3), přičemž součástí prvních dvou je i max pooling. Dále jsou umístěny 3 plně propojené vrstvy, dimenze výstupu poslední vrstvy je pak rovna počtu klasifikovatelných tříd. Během trénování byl navíc mezi konvoluční bloky zapojen i 20% dropout. Celkem síť obsahuje 3 695 202 parametrů.



Obrázek 1: Architektura neurálního modelu

¹ student navazujícího studijního programu Kybernetika a řídicí technika, specializace Umělá inteligence a automatizace, e-mail: simio@students.zcu.cz

4 Tvorba datasetu

K nahrávání, které proběhlo na varně pivovaru Velké Popovice, byl využit mikropočítač s čtyřmi různými mikrofony (celkem 6 kanálů) a pro každou úlohu bylo zvlášť nahráno 16 a 19 hodin dat. Data byla následně dekomponována na jednotlivé kanály a převedena na spektrogramy. Z těchto dat byly pro každou úlohu vytvořeny 2 datasety s rozdílnou délkou okna (5 sekund a 1 sekunda). Zdrojem anotací byl textový log řídicího systému.

Procentuální zastoupení tříd čerpání je 2,2 %, 4,4 % a 3,1 %, zbytek je hluk pozadí. Vysypání se pak vyskytuje v 12,4 % příslušného datasetu, zbytek je opět hluk pozadí. Tato čísla odpovídají reálnému provozu. Před zahájením trénování byla rezervována testovací množina o velikosti 20 % celého datasetu; poměry tříd v obou množinách zůstaly zachované.

5 Parametry trénování a výsledky

Pro každou úlohu byl model trénován samostatně na datasetu příslušnému úloze a samostatně pro každou délku okna. Všechny modely byly trénovány 10 epoch s využitím Adam optimizéru. Konstanta učení byla nastavena na hodnotu 0,001 a jako ztrátová funkce byla zvolena křížová entropie. Výsledky byly vyhodnoceny skrze 3 základní metriky a jsou kompletně zanesené v tabulce 1.

	Klasifikovaná třída	Délka okna: 5 sekund			Délka okna: 1 sekunda		
		Precision	Recall	F1 score	Precision	Recall	F1 score
Čerpání	pozadí	0,9946	0,9921	0,9933	0,9975	0,9940	0,9958
	Čerpání do VK 1	0,7821	0,9396	0,8537	0,8145	0,9597	0,8811
	Čerpání do VK 2	0,9424	0,9784	0,9601	0,9560	0,9801	0,9679
	Čerpání do VK 3	0,9676	0,8483	0,9040	0,9765	0,9191	0,9469
Sypání	pozadí	0,9833	0,9898	0,9865	0,9141	0,9887	0,9500
	Vysypání	0,9169	0,8705	0,8931	0,8129	0,3460	0,4854

Tabulka 1: Evaluace natrénovaných modelů

6 Zhodnocení

F1 skóre jednotlivých tříd ukazuje, že u čerpání dosahuje jednosekundové okno mírně lepších výsledků než okno pětisekundové. Akusticky jsou čerpací režimy charakteristické stabilní amplitudou a frekvencí, kratší okno se tak ukazuje jako dostatečné. Naopak u vysypání sladu se přesnost klasifikace při jednosekundovém okně oproti pětisekundovému výrazně snížila, pravděpodobně kvůli kolísavým akustickým charakteristikám procesu, které závisí např. na úrovni naplnění zásobníku. Tyto charakteristiky dokáže lépe zachytit okno pětisekundové.

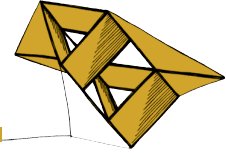
Poděkování

Práce byla podpořena grantem Západočeské univerzity, č. projektu SGS-2025-011.

Literatura

Nandi, P. (2021) *CNNs for Audio Classification*. [Online]

Dostupné z: <https://towardsdatascience.com/cnns-for-audio-classification-6244954665ab/>
[Přístup získán 16 května 2025].



Cross-lingual Emotion Detection

Jakub Šmíd¹

1 Introduction

While emotion detection is well-studied in English, many real-world applications require analysis in languages with limited annotated data. Manual annotation for each language is costly and time-consuming. Cross-lingual settings address this by training models on a high-resource source language (e.g. English) and applying them to target languages without additional supervision. Despite its practical importance, cross-lingual emotion detection remains underexplored. An added challenge is identifying the specific words that express the emotion.

This work presents our approach for the WASSA-2024 *Cross-lingual Emotion Detection Shared Task* (Maladry et al., 2024). The task consists of two subtasks, illustrated in Figure 1: 1) **Cross-lingual emotion detection**: Predicting one of six emotions (*Love, Joy, Fear, Anger, Sadness, Neutral*) for each input sentence; 2) **Trigger word detection**: Identifying which words express the emotion, either in binary format (assigning 0/1 to each token) or numeric format (assigning intensity scores to each token). The task provides English training data, while the evaluation data includes English, Dutch, Russian, Spanish, and French.

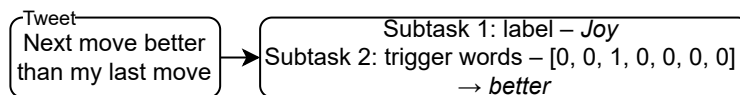


Figure 1: Example tweet with labels for both subtasks.

2 Proposed Solution

We tackle the task using machine translation and Transformer-based models. For the first subtask, we translate the data and fine-tune a large language model (Orca 2 13B).

The second subtask poses additional challenges due to word misalignment between the original and translated texts – such as differing word counts, word order, or multi-word translations. To address this issue, we propose a novel method. We mark each trigger word in English with special symbols (e.g. “[]” or “{ }”) before translation and then recover the trigger words from the translated text using these markers, as shown in Figure 2.

We further enrich the training data by creating bilingual sentence pairs: one in the source language with trigger words translated and another in the target language with trigger words retained in English, as illustrated in the lower part of Figure 2. This approach yields four training sets: the original English data, its translations into four target languages, and two trigger-word-switched sets. We fine-tune two multilingual models – mT5 and XLM-R – on all four sets. Including the trigger-word-switched data significantly improves performance, confirming the effectiveness of our approach.

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

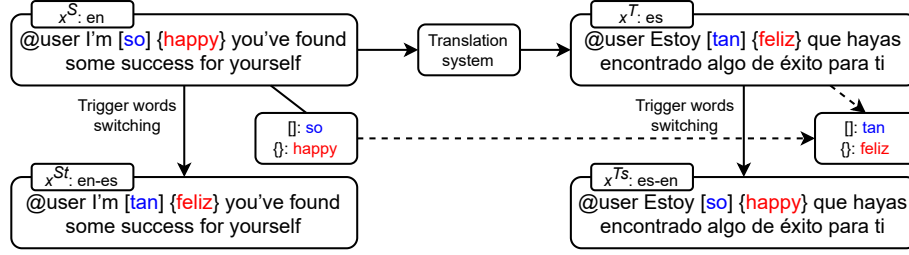


Figure 2: Example of label projection method with trigger word switching (lower part) for English and Spanish language pair.

3 Achieved Results

Table 1 shows our results on all three subtasks, our rank and the best system performance in the competition. Our system ranked **first** in numerical trigger word detection, **third** in binary trigger word detection, and **seventh** in cross-lingual emotion classification.

Subtask	Result	Rank	Best
Cross-lingual emotion detection	59.10	7.	62.95
Trigger word detection (binary)	59.19	3.	61.58
Trigger word detection (numerical)	70.52	1.	70.52

Table 1: Results for all subtasks, including system rank and best score in the competition.

4 Conclusion

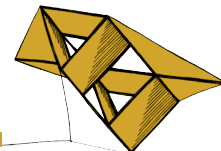
We propose a method combining machine translation with large and multilingual Transformer-based models to address emotion classification and trigger word detection in cross-lingual settings. For emotion classification, we fine-tune Orca 2 13B. We introduce a novel approach for trigger word detection involving trigger word marking, label projection, and training data augmentation with bilingual sentence pairs featuring trigger word switching. We use this enriched data to tune the multilingual Transformer models. Our system achieves excellent results in the competition, ranking **1st** in numerical trigger word detection, **3rd** in binary detection, and **7th** in cross-lingual emotion classification, validating the effectiveness of our approach.

Acknowledgment

This work has been partly supported by the Grant No. SGS-2022-016 Advanced methods of data processing and analysis and by the OP JAC project DigiTech no. CZ.02.01.01/00/23_021/0008436. 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

Aaron Maladry, Pranaydeep Singh, and Els Lefever. 2024. Findings of the WASSA 2024 EXALT shared task on Explainability for Cross-Lingual Emotion in Tweets. In *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 454–463, Bangkok, Thailand. Association for Computational Linguistics.



Využití transformerů pro specifickou úlohu z praxe

Daniel Tauš¹

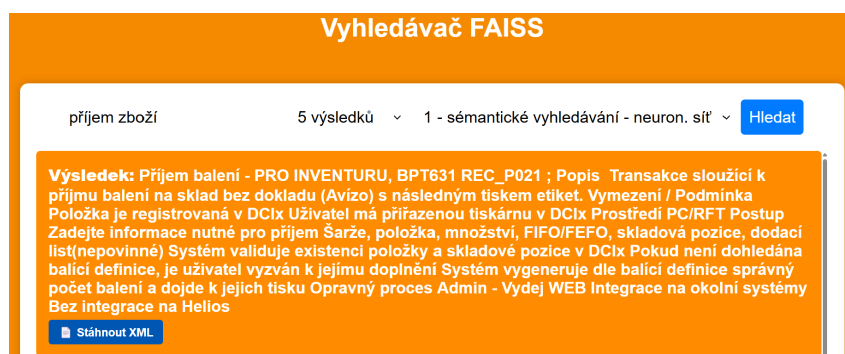
1 Úvod

Moderní podnikové systémy, například v logistice, často využívají transakční modely reprezentované XML. Manuální tvorba dokumentace k těmto transakcím a její následné vyhledávání jsou však časově náročné a vedou k nekonzistencím, což omezuje efektivitu. Tato práce proto zkoumá využití modelů architektury Transformer, představenou v práci Vaswani et al. (2017), ke zlepšení těchto procesů. Konkrétně se zaměřuje na dvě hlavní oblasti:

- automatické generování textových popisů transakcí z jejich XML dat;
- vývoj systému pro efektivní sémantické a hybridní vyhledávání v těchto popisech.

2 Sémantické a Hybridní Vyhledávání Transakcí

Klíčovým úspěchem práce je vývoj a evaluace systému pro efektivní vyhledávání transakčních popisů na základě uživatelského dotazu. Jádrem tvoří sémantické vyhledávání s modelem Jina Embeddings v3, který představil Sturua et al. (2024) (pro texty až 8192 tokenů) a index Faiss. Tento model byl vybrán po srovnání, kde dosáhl 100% úspěšnosti na testovací sadě a podporoval dlouhé vstupy. Dimenze embeddingů 768 poskytla nejlepší relevanci. Vyhledávání funguje na základě významu. Pro zlepšení relevance byl systém rozšířen o hybridní přístup kombinující sémantické skóre (Faiss) s TF-IDF. Uživatel váží poměr mezi těmito přístupy. Efektivní dvoufázový výpočet umožňuje provoz na běžném kancelářském počítači. Výsledný prototyp ve formě webové aplikace byl interně pozitivně přijat a jeho upravená verze implementována ve firmě, což potvrzuje jeho praktickou použitelnost.

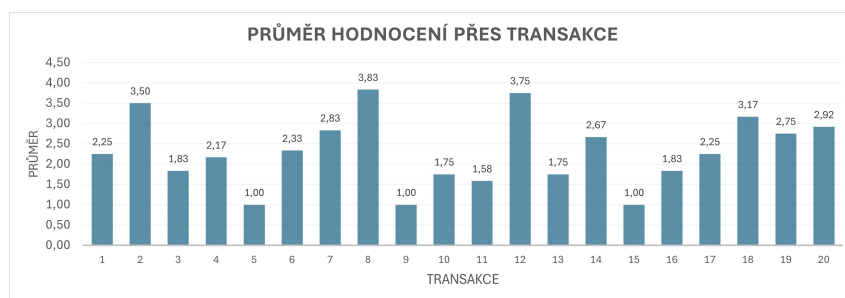


Obrázek 1: Hybridní sémantický vyhledávač

¹ student navazujícího studijního programu Kybernetika a řídicí technika, specializace Umělá inteligence a automatizace, e-mail: dantauš6@students.zcu.cz

3 Generování Popisů Transakcí

Práce experimentovala s generováním popisů transakcí z XML modelem T5, představeným v práci Raffel et al. (2019), pro automatizaci dokumentace, s testováním různých formátů vstupních dat a úrovní detailů. Výsledky byly smíšené: T5 generoval korektní texty pro jednoduché transakce, zachovával strukturu. U komplexnějších ale vynechával detaily, chyboval v interpretaci (např. SQL) či tvořil nekonzistence. Kvalitu ovlivnila bohatost vstupů. Uživatelský průzkum (12 expertů hodnotilo 20 popisů na škále 1-5) s průměrným skóre 2.31 (medián 2) potvrdil, že popisy nejsou připraveny k nasazení bez manuálních úprav; více než 61% vyžadovalo úpravy. Hlubší analýza těchto hodnocení, vizualizovaná na Obr. 2, odhalila zásadní zjištění v podobě výrazné heterogenity v úspěšnosti generování napříč různými typy transakčních procesů. Průměrná hodnocení pro jednotlivé transakce totiž vykazovala značný rozptyl, pohybující se v intervalu od 1.00 do 3.83. Tato pozorovaná variabilita zdůrazňuje kritickou závislost kvality generovaných textů na komplexitě a specifikách vstupních dat, přičemž model T5 projevoval nedostatky zejména u komplexnějších transakcí. Inverzní úloha (generování struktury z textu pomocí T5) nebyla úspěšná díky nelogickým výstupům, zejména kvůli halucinování modelu a velmi složité požadované struktuře dat. Do budoucna by úlohu generování popisů transakcí, zejména s ohledem na komplexitu a požadovanou kvalitu, mohlo výrazně vylepšit nasazení moderních velkých jazykových modelů (LLM), které disponují pokročilejšími schopnostmi porozumění kontextu a generování přirozeného jazyka.



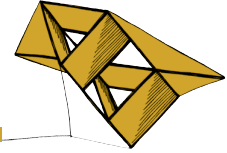
Obrázek 2: Průměrné hodnocení kvality popisů pro jednotlivé transakce

Poděkování

Výpočetní prostředky byly poskytnuty projektem e-INFRA CZ (ID:90254), podporovaným Ministerstvem školství, mládeže a tělovýchovy České republiky.

Literatura

- A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin (2017). Attention is all you need. *Advances in neural information processing systems*.
- C. Raffel, N. Shrivastava, Z. Li, A. Matena, M. McCoy, B. Bast, N. Shazeer, W. Chiang, S. Chen, B. Doherty, et al. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv preprint arXiv:1910.10683*.
- S. Sturua, I. Mohr, M. K. Akram, M. Günther, B. Wang, M. Krimmel, F. Wang, G. Mastrapas, A. Koukounas, N. Wang, and H. Xiao (2024). jina-embeddings-v3: Multilingual Embeddings With Task LoRA. *arXiv:2409.10173 [cs.CL]*. <https://arxiv.org/abs/2409.10173>.



Point-mass Filter with Non-equidistant Grid Design

Jan Trejbal¹

1 Introduction

State estimation is essential in engineering applications from navigation to control. While many estimation methods assume Gaussian probability density functions (PDFs), global filters can capture more complex, non-Gaussian PDFs. The point-mass filter (PMF) is a global filter that discretises the state-space into a grid and approximates the PDF as a piecewise constant function, known as the point-mass density (PMD). Unlike particle filters, the PMF produces deterministic estimates: given the same measurements, it will always generate identical outputs. This deterministic nature, combined with its structured representation of the entire distribution, higher robustness and better handling of abrupt changes, makes it valuable in high-reliability applications like navigation systems, see Anonsen and Hallingstad (2006).

2 Point-mass Filter and Motivation

The grid layout is recognised as a crucial aspect of the PMF, since it heavily influences the performance. A sufficiently large portion of the state-space must be covered to capture the essential mass of the conditional PDF, while keeping computational demands reasonable. A common strategy is to estimate the first two predictive moments—mean $\hat{x}_{k+1|k}$ and covariance $C_{k+1|k}$ —and use them to define the region over which the grid is placed. In each state dimension, this region typically spans the interval $\pm L(i) = \eta \sqrt{C_{k+1|k}(i, i)}$, where η is a user-defined scaling parameter. The full grid is usually constructed as a Cartesian product of vectors defined in each dimension, resulting in an equidistant grid.

In this work, we explore whether a non-equidistant placement of grid points within the same predictive region can improve performance for a fixed number of points. The goal is to allocate points more efficiently by concentrating them in regions of expected higher probability, potentially enhancing estimation accuracy without using more points.

3 Non-equidistant Grid Construction

Proposed method deals with creating vector of $N(i)$ points in each dimension. Rather than defining point locations directly, $N(i) + 1$ borders are first calculated, to ensure proper coverage of the area, between $-L(i)$ and $L(i) \forall i$, and the position of points and their neighbourhoods is derived from the positions of borders between them. To achieve this, a symmetric exponential transformation was employed:

1. A vector s is calculated - for odd $N(i)$ it is calculated as $s = ((1 : m) - 0.5)/(m - 0.5)$;, where $m = (N(i) + 1)/2$
2. The vector s is then transformed to receive the positive borders

$$s \mapsto b(s) = L(i) \frac{e^{\alpha s} - 1}{e^{\alpha} - 1}$$

¹ student of the doctoral degree program Applied Sciences and Informatics, field of study Cybernetics, e-mail: trejbalj@students.zcu.cz

3. Positive borders are mirrored around zero to receive negative borders; together, they define the vector of points.

For $\alpha > 0$, more points are put near the centre of the grid, for $\alpha < 0$, fewer points are put near the centre of the grid, and for $\alpha \rightarrow 0$, points are distributed equidistantly, as in standard equidistant grid, as shown in Figure 1.

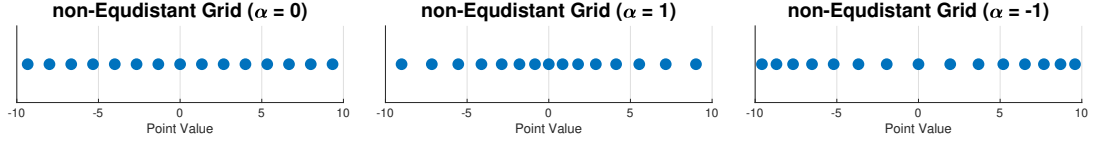


Figure 1: Comparison of 1D grids with different α values

4 Simulation Results

The grid creation rule based on the design parameter α influences the shape of the predictive grid and consequently affects the approximation of the underlying PDF. To evaluate this influence, we conducted a series of Monte Carlo simulations with different types of initial PDFs and system noise models. For each scenario, we began with a common initial grid and generated multiple predictive grids using different values of α . We then computed the resulting predictive and filtering PMDs and analysed how variations in α affected various performance metrics. During the simulation, the same values of α were used in all dimensions. An example of the results is shown in Figure 2.

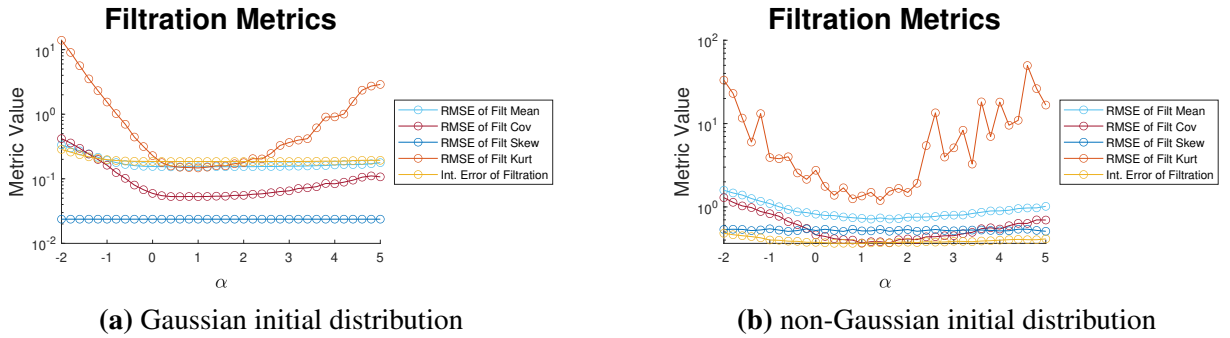


Figure 2: Part of simulation results

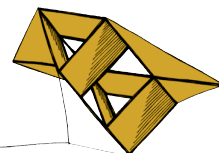
Based on the results, we conclude that an appropriate selection of α can significantly improve performance outcomes. The optimal choice of α depends on three key factors: the specific metric we aim to minimise, the underlying system model characteristics, and the shape of the underlying true PDF. Identifying suitable adaptation strategies for α will be the subject of future research.

Acknowledgement

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

References

- Anonsen, K., and Hallingstad, O. (2006) Terrain-aided underwater navigation using point mass and particle filters. *Proceedings, IEEE/ION Position, Location, and Navigation Symposium*. Coronado, CA, USA, pp. 1027–1035.



Nonsense Word Repetition for Detection of Cognitive Disorders

Jan Tupý¹

1 Introduction

Cognitive impairments are neurological disorders that disrupt fundamental mental functions such as memory, attention, or language. Although they most commonly appear in older adults, they can also result from brain injuries or psychiatric conditions. Due to the increasing prevalence of these disorders, there is a growing emphasis on early and reliable diagnosis. Traditional diagnostic procedures are time-consuming, which has led to a rising interest in automated digital testing.

One innovative approach involves testing with **nonsense words** that resemble regular vocabulary but lack meaning. This method enables researchers to focus on pure cognitive processes without interference from semantic context. The aim of this study is to investigate whether this type of task can provide objective and quantitative data on the state of cognitive functions and support the development of effective digital diagnostic tools.

1.1 Nonsense Words

The nonsense words were selected by a cognitive disorders expert, with an emphasis on phonetic naturalness and the absence of semantic meaning. These words are part of one of ten tasks included in a broader electronic test battery, which is currently undergoing pilot testing. A key advantage of this digital format is that it allows participants to complete the test independently from home using a personal computer, without the need for professional supervision. The following words are used in our work:

čapiváček, kokávka, flonka, lampešiška, trebice, cidla

2 Methodology

Participants were sequentially presented with six nonsense words (see Section 1.1), pre-recorded using a human voice as reference templates. At the beginning of the task, participants were instructed to immediately repeat each word aloud after it was played. This procedure continued until all six words were processed.

Speech was recognized using a model based on the modern **wav2vec 2.0** architecture, specifically the system described in Lehečka et al. (2022). The recognition pipeline employs a grapheme-based approach complemented by an external language model. The recognized output was preprocessed by removing punctuation and converting all text to lowercase. As the final step, the output was compared with the reference words using strict matching.

The dataset consisted of 149 participants, including 108 healthy controls and 41 individuals diagnosed with cognitive impairment. For binary classification (0 = healthy, 1 = patient), LogisticRegression, RandomForest, XGBoost, and HistGradientBoosting models were

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

employed, implemented using the `scikit-learn` and `xgboost` libraries. Model optimization focused primarily on the `F1-score`, while also monitoring other metrics such as `accuracy`, `recall`, and `precision`. Evaluation was conducted using five-fold stratified cross-validation.

3 Results and Analysis

An analysis was conducted on the success rate of repeating individual nonsense words in healthy participants and those with cognitive impairment. A comparison of strict match results between the two groups is shown in Figure 1, illustrating the average accuracy for each word, with shaded areas representing 95% confidence intervals.

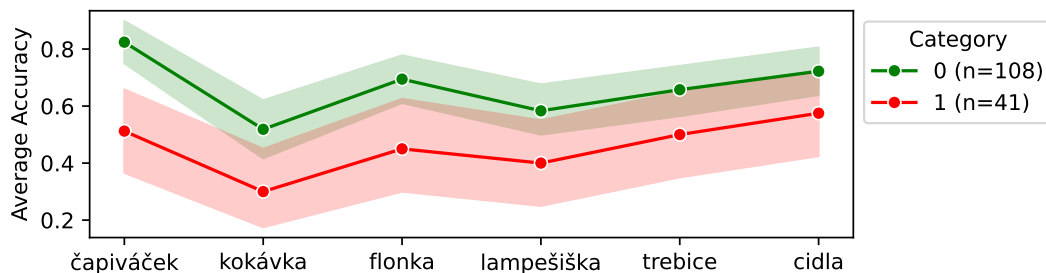


Figure 1: Average Strict Match Accuracy (0 to 1)

The results indicate that the `Logistic Regression` model achieved the highest `F1-score`, suggesting its better ability to distinguish between healthy participants and those with cognitive impairment. A complete overview of the metrics for each model is provided in Table 1.

Table 1: Classification results for individual models.

Model	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.607	0.632	0.599	0.631
Random Forest	0.557	0.555	0.556	0.651
Hist. Grad. Boosting	0.610	0.556	0.550	0.718
XGBoost	0.677	0.590	0.593	0.745

4 Conclusion

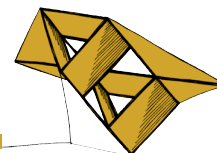
The results show that the task of repeating nonsense words reveals differences between healthy individuals and patients with cognitive impairment, although the current predictive accuracy is not yet suitable for clinical use. Despite limited performance, the data suggest potential for this method, particularly when combined with other tasks or additional features beyond strict match. In this sense, it represents a promising direction for the future development of accessible, digitally implementable screening tools.

Acknowledgement

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

References

Lehečka, J., Švec, J., Pražák, A., Psutka, J. (2022) *Exploring Capabilities of Monolingual Audio Transformers using Large Datasets in Automatic Speech Recognition of Czech*. In: *Proc. Interspeech 2022*, pp. 1831–1835.

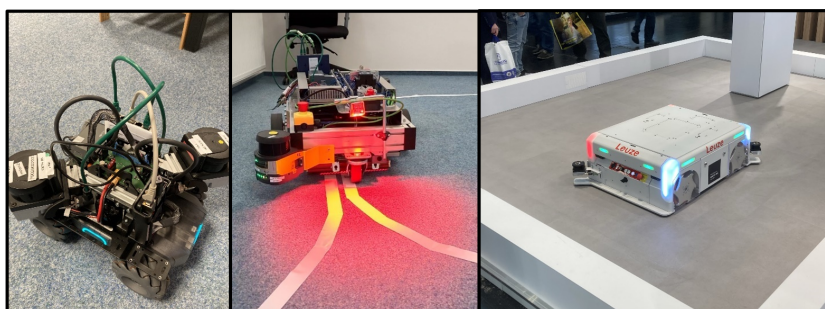


Vývoj autonomní mobilní platformy pro přepravu nákladů

Jakub Tvrz¹

1 Úvod

Autonomní mobilní roboty (AGV a AMR) jsou významným prvkem průmyslové automatizace. Tento příspěvek představuje otevřený a modulární řídicí systém založený na frameworku ROS2 [1], který překonává omezení uzavřených systémů, jako je SIMOVE, a nabízí pokročilé navigační přístupy, detekci palet z 2D LiDARů a bezpečný provoz podle norem SIL. Hlavním přínosem je adaptabilita systému na různé podvozky a modularita díky dockerizaci jednotlivých funkčních balíčků, což zajišťuje snadné nasazení, přenositelnost a škálovatelnost. Vyvinutý driver zajišťuje komunikaci mezi PLC a ROS2, zatímco webová aplikace (vyvinutá kolegou J. Holubem [2]) umožňuje intuitivní ovládání. Systém podporuje integraci s existujícím hardwarem, například při repasi robotů. Práce poskytuje komplexní vhled do problematiky a návod, jak vybudovat kompletní řídicí systém průmyslového AMR.



Obrázek 1: a) DJI model b) Diferenciální AGV c) Velké všesměrové AGV

Systém byl vyvinut a testován na čtyřech podvozcích: **DJI** sloužil pro počáteční ověření algoritmů na reálném robotu a k integraci senzorů, **diferenciální podvozek** umožnil první testování safety chování splňujícího bezpečnostní normy a testování komplexních scénářů v reálných skladech při použití navigace po čáře, **velké AGV s všesměrovým podvozkem** je finálním výstupem a je zde nasazena plně autonomní navigace i detekce palet, diferenciální podvozek **TinyAGV** byl použit pro vývoj cyklického sledování virtuálních drah, které si uživatel nakreslí do mapy a simulační model v Gazebo poskytl prostředí pro prvotní validaci algoritmů.

2 Navržené přístupy a struktura řídicího systému

Mapování a lokalizace je základem všech vyvinutých přístupů. Mapování prostředí je realizováno pomocí SLAM (Simultaneous Localization and Mapping [3]) algoritmu, který vytváří 2D mapu prostředí na základě LiDARových skenů a odometrických dat. Lokalizace využívá algoritmus AMCL (Adaptive Monte Carlo Localization), který koriguje drift odometrie porovnáním LiDARových skenů s mapou a je založen na částicovém filtru. Autonomní navigace

¹ student navazujícího studia fakulty Aplikovaných věd, obor studia Kybernetika,
e-mail: tvrzjak@students.zcu.cz

navrženého systému kombinuje globální a lokální plánování trajektorií pro efektivní a bezpečný pohyb vozíku. Globální plánování, realizované algoritmem NavFn, generuje optimální trasu na základě statické nákladové mapy a cílové pozice určené uživatelem, s ohledem na statické překážky. Lokální plánování, založené na MPPI (Model Predictive Path Integral Control), dynamicky upravuje pohyb v reálném čase podle aktuálních LiDARových dat a odometrie v oblasti kolem robota a přizpůsobuje trajektorii dynamickým překážkám. MPPI generuje a hodnotí řadu možných cest podle nákladové funkce (vzdálenost, překážky, dynamika) a vybírá optimální cestu v rámci plovoucího horizontu, zatímco vyhlazovač zajišťuje plynulost pohybu minimalizací zrychlení a respektováním kinematických omezení robota.

Důležitou součástí AGV je schopnost detekovat a nakládat palety. Standardně jsou k tomu používány kamery a rozpoznávání obrazu. V této práci byl navržen inovativní přístup založený výhradně na datech z 2D LIDARových senzorů, který nabízí ekonomickou výhodu již integrovaných LIDARů, robustnost vůči změnám prostředí a osvětlení a vysokou rychlost zpracování. Detekce objektů z LIDARových skenů v reálném čase využívá shlukování, konkrétně algoritmus DBSCAN. Z těchto objektů je možné dále detekovat podle vzájemné orientace a polohy složitější tvary v prostoru, jako je například paleta. Využití však může být i v trekování dynamických objektů a predikci kolizí.

Dalším přístupem jsou virtuální dráhy, jejichž uživatelská přívětivost spočívá v tom, že uživatel může nakreslit uzavřenou křivku do mapy a robot ji bude automaticky sledovat. Implementace vychází z vytvoření mapy prostředí, která slouží jako základ pro definici trajektorie. Následně je aktivována lokalizace robota v mapě. Fixní trajektorie, definovaná jako posloupnost bodů, je sledována v cyklickém režimu algoritmem Pure Pursuit, který zajišťuje plynulé přizpůsobení směru a rychlosti podle aktuální polohy a orientace robota.

Trochu stranou stojí návrh standardního systému jízdy po čáře s čtením QR kódů. Je založený na využití průmyslových senzorů a PD regulátoru s více rychlostními stupni a bezrázovým přepínáním. Během jízdy nejsou opominuty bezpečnostní funkce zajištěné LIDARy.



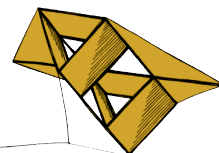
Obrázek 2: a) *TinyAGV a virtuální dráha*, b) *Simulace a ovládací prostředí RViz*

Poděkování

Příspěvek vychází z diplomové práce, kterou zpracovávám pod záštitou firmy Leuze Engineering Czech, s.r.o.

Literatura

- [1] ROS 2 Documentation 2014; ROS 2 Documentation: Humble documentation — docs.ros.org. <https://docs.ros.org/en/humble/index.html>. [Accessed 05-03-2025].
- [2] Jan Holub. Development of control system and handling interface for automated guided vehicle. FEL ČVUT, 05 2025.
- [3] Steve Macenski and Ivona Jambrecic. Slam toolbox: Slam for the dynamic world. *Journal of Open Source Software*, 6(61):2783, 2021.



Using Pre-trained Models for Phoneme Representation in Czech Speech Synthesis

Lukáš Vladař¹

1 Introduction

Text-to-speech (TTS) systems, i.e., systems producing artificial speech, represent an important topic in the field of artificial intelligence. Modern approaches based on neural networks reach very good results, almost comparable to real human speech.

Nguyen et al. (2023) argue that including a large-scale pre-trained model for phoneme representation in a neural TTS system can further improve the final synthetic speech. We used their pre-trained model called *XPhoneBERT* to investigate whether it can also enhance the quality of speech synthesis in the Czech language.

2 VITS and XPhoneBERT

For our experiments, we used *VITS*, which is a widely used end-to-end TTS model introduced by Kim et al. (2021). *VITS* contains a text encoder, which processes the input phonetic representation of the synthesized utterance.

Instead of the default encoder, Nguyen et al. (2023) used their own model called *XPhoneBERT*, which is based on the *BERT* model described by Devlin et al. (2019). *XPhoneBERT* has been trained on a huge multi-lingual dataset to learn 192-dimensional phoneme representations. The authors claim that using *XPhoneBERT* as a text encoder of the *VITS* model improves the naturalness and prosody of the output synthetic speech.

3 Experiment Description

We performed an experiment to see whether the use of the *XPhoneBERT* encoder can also raise the quality of Czech speech synthesis. We trained two original *VITS* models, each of them was trained using 18 hours of speech recorded by a professional Czech speaker. One of the speakers was male and the other female. Furthermore, we used the same datasets to train *VITS* synthesizers combined with the pre-trained *XPhoneBERT* encoder.

Subsequently, we compared the original *VITS* models and the models with the *XPhoneBERT* encoder in a preference listening test. The test was performed by five listeners, all of whom were Czech native speakers. Throughout the test, 30 utterances were presented to the listeners. Each utterance was represented by two recordings, and the listeners were required to choose the better-sounding version. If they were not able to choose the better-sounding recording, they were allowed to state that the quality of both is approximately the same.

When the listening test was finished, we counted how many times the original *VITS* model was selected to sound better, and how many times the combination with *XPhoneBERT* was preferred. We used these numbers to perform the one-sided sign test, which is described in

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

Conover (1999), in order to determine the statistical significance of the results.

4 Results

Figure 1 shows the distribution of the answers obtained in the listening test. Both models achieved very similar ratings when trained on the female speaker dataset. However, VITS combined with the XPhoneBERT encoder reached a significantly better result in the case of the male voice, which was also proved by a one-sided sign test.

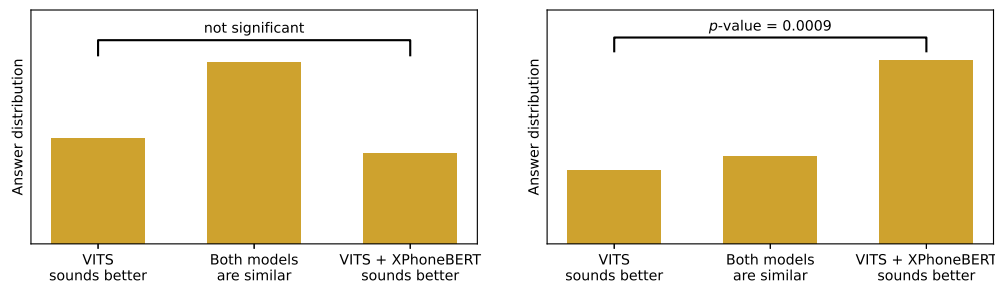


Figure 1: The distribution of different answers obtained in the listening test comparing models trained using female (on the left) and male (on the right) speaker’s recordings

5 Conclusion

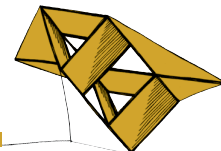
We showed that including a pre-trained text encoder in a TTS model may improve the quality of Czech speech synthesis. However, it may not help in some cases. The XPhoneBERT encoder, which we used for our experiments, has been trained on a huge multi-lingual dataset, however, this dataset contained only 4542 Czech sentences, while US English, for example, was represented by 33515 sentences (see Nguyen et al. (2023)). This imbalance may be the reason why XPhoneBERT does not improve speech synthesis in Czech as much as in other languages.

Acknowledgement

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

- Conover, W.J. (1999) ‘The Sign Test’ in *Practical Nonparametric Statistics, 3rd Edition*. New York: John Wiley & Sons, pp. 157-165.
- Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. (2019) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *Proceedings of NAACL*. Minneapolis, Minnesota, June 2021. Association for Computational Linguistics. pp. 4171-4186.
- Kim, J., Kong J., Son, J. (2021) Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech, *Proceedings of the 38th International Conference on Machine Learning*. July 18-24, 2021. PMLR. pp. 5530-5540.
- Nguyen, L.T., Pham T., Nguyen, D.Q. (2023) XPhoneBERT: A Pre-trained Multilingual Model for Phoneme Representations for Text-to-Speech, *Proc. Interspeech 2023*. Dublin, August 20-24, 2023. ISCA. pp. 5506-5510.



Towards Aesthetic Enrichment of Mirror Selfies via Automatic Image Analysis

Jiří Vyskočil¹, Kristina Zejkanová²

1 Introduction

Mirror selfies have become a common form of self-expression on social media, often shared for their aesthetic value. This trend inspired a collaboration with the Ladislav Sutnar Faculty of Design and Art to explore how design elements — such as chairs or decorative objects — might enhance these images. We developed a system for automatic analysis of mirror selfies collected from social networks. The goal is to automatically extract visual metadata, including color palettes, human poses, face occlusion by phones, subject placement (e.g., golden ratio), and object directions relative to the person. The obtained metadata can support the creation of design elements that make selfies more visually appealing.

2 Methodology

To solve the task, we used three pre-trained models: OneFormer by Jain et al. (2023) for image segmentation, YOLOv8 by Jocher et al. (2023) for pose estimation, and InsightFace by Deng et al. (2020) for face detection (see Figure 1). We kept only relevant segmentation classes (e.g., person, phone, bed, chair) and selected the person with the most keypoints as the main subject.

Face visibility was computed as the intersection between the face and phone masks relative to the face area. Color palettes were extracted using the k-means and mean-shift algorithms in the Hue-Saturation-Value (HSV) color space for improved accuracy. Pose classification (standing vs. sitting) was based on vector direction and magnitude heuristics derived from keypoint positions; cases with missing leg information were labeled as unknown.

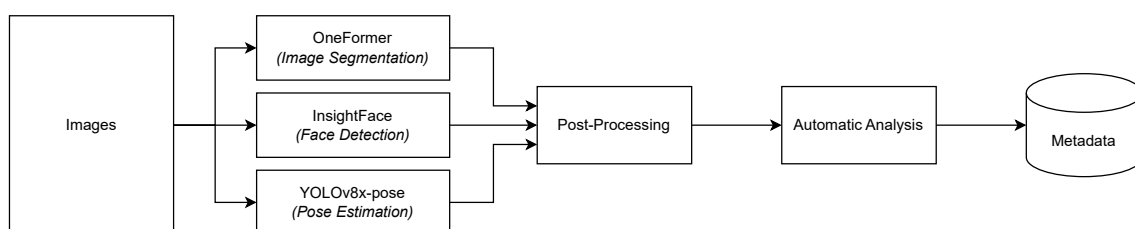


Figure 1: Overview of the proposed system. Input images are processed by deep learning models, followed by post-processing (e.g., filtering relevant COCO categories, approximating face masks as ellipses). Finally, the measurements are transformed into the desired metadata.

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

² PhD student, Interdisciplinary Research through Visual Arts program, e-mail: kzejcano@students.zcu.cz

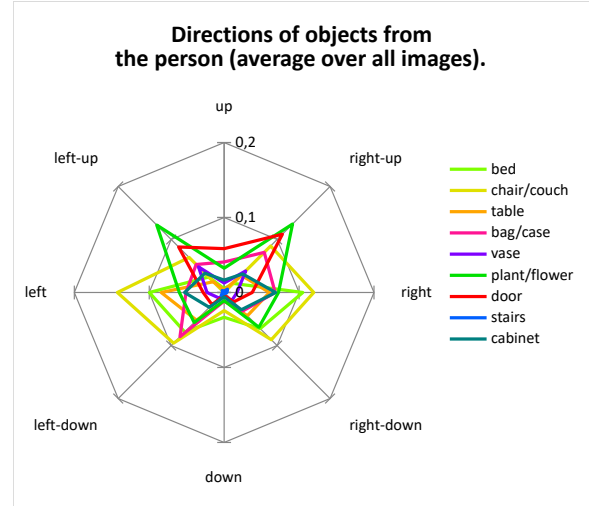
3 Results and Conclusion

The analysis of available mirror selfies revealed a frequent use of light blue clothing, mostly jeans, often paired with other light tones such as cashmere or light grey. In most images, the face is partially or fully covered by the phone, and subjects are typically standing (see Figure 2 (a)). On average, the person occupies 11.26% of the total image area. Photos were most often taken in bedrooms or living rooms, and less frequently in hallways. This is reflected in the object direction visualization (Figure 2 (b)).

These findings will inform the next phase — designing aesthetic elements, such as elevated chairs, that align with the observed styles and color palettes to enhance the visual appeal of mirror selfies.

Category	Label	Count
Face visibility	Visible	607
	Not visible	965
Pose	Sitting	210
	Standing	1322
	Unknown	40

(a) Results of face visibility and pose classification. Visibility was determined by overlap between face and phone masks (threshold 50%), and pose was inferred using a YOLO-based detector with heuristic rules.



(b) Visualization of the relative positions of objects from the centroid of the person, showing their directional orientation.

Figure 2: Overview of the automatic annotation results obtained from 1572 images. (a) Distribution of face visibility and person pose; (b) Relative positions of objects from the person.

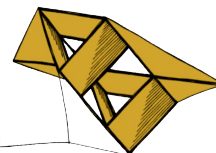
4 References

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:90140), supported by the Ministry of Education, Youth and Sports of the Czechia. We would like to thank ChatGPT for its help with language improvements.

References

- Jocher, G., Qiu, J., Chaurasia, A. (2023) *Ultralytics YOLO (Version 8.0.0)* [Software]. Available from: <https://github.com/ultralytics/ultralytics> [Accessed May 2025].
- Jain, J., Li, J., Chiu, M.T., Hassani, A., Orlov, N., Shi, H. (2023) OneFormer: One Transformer to Rule Universal Image Segmentation. *Proceedings of the IEEE/CVF CVPR*.
- Deng, J., Guo, J., Ververas, E., Kotsia, I., Zafeiriou, S. (2020) RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild. *Proceedings of the IEEE/CVF CVPR*.



Complexity of KIR allele identification pipeline design

Kateřina Wolf^{*1}, Filip Jani¹, Daniel Georgiev¹, Pavel Jindra^{3,4}, Monika Holubová^{2,3} and Lucie Houdová¹

1 Introduction

In Czechia, 1200 people are diagnosed with hemato-oncological disease each year. Treatment of this kind disease involves hematopoietic stem cell transplantation (HSCT), and its' success depends on choosing the suitable donor. Although genetic variant in genes other than Human Leukocyte Antigen (HLA) are not primary criteria for the donor selection, their impact can play a role in the patient outcome and it is usually involved in post-transplant conditions (GVL, GVHD). Killer-cell immunoglobulin-like receptor (KIR) genes, affecting HSCT outcome by presence/absence of specific gene, are currently studied for their potential through gene variants (alleles). An analytical pipeline that enables quick and easy assessment of KIR allele is desirable. However, during the pipeline design, it is necessary to address and manage certain issues from both data science and medical point of view.

2 Materials and methods

For the development of an analytical method for KIR identification, we used synthetic data based on reference sequences IPD ([3], 2.11.0). Genotype sensitivity was examined on synthetic datasets corresponding to actual genotypes [5]. Pipeline setting is limited only to known alleles. Validation was based on well-characterized cell lines [4].

3 Results

During the analysis of data and issues for pipeline creation, the identified problems were addressed and managed. Our approach reaches the lowest accuracy in the 2DL1 gene, which was successfully identified in 73% of cases, while the other genes were successfully identified in more than 88% of cases. The most precise results are for the 2DL4 gene, which was correctly identified in 99% of cases. In the case of real data, the results of identification were lower as expected due to biological factors.

4 Conclusion

The design and implementation of the KIR allele identification pipeline is a complex task with potential use in clinical practice. The correctness and precision of the pipeline are influenced not only by the technical design itself but also by input data and the chosen reference sequence.

^{*} Phd student, Applied Sciences, field of study cybernetics, e-mail: kkratoch@ntis.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

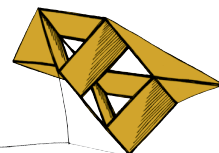
⁴ Czech National Marrow Donors Registry, Czechia

Acknowledgement

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

References

- ¹ Pende D, Falco M, Vitale M, Cantoni C, Vitale C, Munari E, Bertaina A, Moretta F, Del Zotto G, Pietra G, Mingari MC, Locatelli F and Moretta L (2019) Killer Ig-Like Receptors (KIRs): Their Role in NK Cell Modulation and Developments Leading to Their Clinical Exploitation. *Front. Immunol.* 10:1179. doi: 10.3389/fimmu.2019.01179
- ² Marin WM, Dandekar R, Augusto DG, Yusufali T, Heyn B, Hofmann J, et al. (2021) Highthroughput Interpretation of Killer-cell Immunoglobulin-like Receptor Short-read Sequencing Data with PING. *PLoS Comput Biol* 17(8): e1008904 . doi: 10.1371/journal.pcbi.1008904
- ³ James Robinson, Jason A. Halliwell, Hamish McWilliam, Rodrigo Lopez, Steven G. E. Marsh, IPD—the Immuno Polymorphism Database, *Nucleic Acids Research*, Volume 41, Issue D1, 1 January 2013, Pages D1234–D1240, doi: 10.1093/nar/gks1140
- ⁴ Maniangou B, Legrand N, Alizadeh M, Guyet U, Willem C, David G, Charpentier E, Walencik A, Retière C and Gagne K (2017) Killer Immunoglobulin-Like Receptor Allele Determination Using Next-Generation Sequencing Technology. *Front. Immunol.* 8:547. doi: 10.3389/fimmu.2017.00547
- ⁵ Gonzalez-Galarza FF, McCabe A, Santos EJMD, Jones J, Takeshita L, Ortega-Rivera ND, Cid-Pavon GMD, Ramsbottom K, Ghattaoraya G, Alfirevic A, Middleton D, Jones AR. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res.* 2020 Jan 8;48(D1):D783-D788. doi: 10.1093/nar/gkz1029. PMID: 31722398; PMCID: PMC7145554.
- ⁶ Wagner I, Schefzyk D, Pruschke J, Schöfl G, Schöne B, Gruber N, Lang K, Hofmann J, Gnahn C, Heyn B, Marin WM, Dandekar R, Hollenbach JA, Schetelig J, Pingel J, Norman PJ, Sauter J, Schmidt AH and Lange V (2018) Allele-Level KIR Genotyping of More Than a Million Samples: Workflow, Algorithm, and Observations. *Front. Immunol.* 9:2843. doi: 10.3389/fimmu.2018.02843



Detekce anomálií v datech z knih limitních objednávek

Dominik Zappe¹

1 Abstrakt

Moderní finanční trhy jsou rychlé, komplexní a stále častěji se stávají cílem sofistikovaných forem manipulace. Tato práce se zaměřuje na detekci anomálií v časových řadách odvozených z knih limitních objednávek s cílem rozpoznat manipulativní chování zvané *spoofing*. Vzhledem k absenci anotovaných dat jsou použity metody strojového učení bez učitele aplikované na reálná historická data. V práci je implementováno šest metod – izolační les, lokální faktor odlehlosti, jednotřídní SVM, plně propojený autoenkodér, konvoluční autoenkodér a transformer autoenkodér. Modely jsou evaluovány pomocí méně známých metrik *Excess Mass* a *Mass Volume*, přičemž nejlépe si vedou modely izolační les a transformer. Kombinací nejvýkonnějších modelů vznikl robustní nástroj schopný odhalit podezřelé chování bez ruční anotace. Navržené řešení efektivně identifikuje rizikové oblasti pro následnou expertní analýzu a představuje tak praktický přínos pro detekci nelegálních praktik na finančních trzích.

2 Úvod

Cílem práce je navrhnout a implementovat rámec pro detekci anomálií v časových řadách z knih limitních objednávek, se zaměřením na odhalení **spoofingu** – manipulativního chování na finančních trzích. Důraz je kladen na použití metod *strojového učení bez učitele* nad neanotovanými daty.

3 Terminologie a metodologie

Limitní kniha objednávek (*Limit Order Book, LOB*) je dynamická databáze, která v reálném čase eviduje příkazy k nákupu a prodeji finančních instrumentů, viz Abergel et al. (2016). Obsahuje **nabídkovou** (*Bid*) a **poptávkovou** (*Ask*) stranu, přičemž objednávky jsou řazeny podle ceny a času zadání.

Spoofing představuje formu tržní manipulace, při níž jsou zadávány rozsáhlé objednávky bez záměru jejich realizace, s cílem vyvolat falešný dojem tržní aktivity. Po dosažení požadovaného cenového pohybu jsou tyto příkazy stornovány, viz Montgomery (2016).

Vzhledem ke krátkodobému charakteru *spoofingu* a absenci anotovaných dat je jeho odhalování obtížné. Práce se proto zaměřuje na jeho detekci prostřednictvím *metod učení bez učitele*.

4 Implementovaná řešení

V rámci práce bylo implementováno šest detekčních algoritmů:

- **Izolační les** (*Isolation Forest*)
- **Lokální faktor odlehlosti** (*Local Outlier Factor*)
- **Jednotřídní SVM** (*One-Class SVM*)

¹ student navazujícího studijního programu Informatika a její specializace, specializace Zpracování přirozeného jazyka, e-mail: zapped99@students.zcu.cz

- **Plně propojený autoenkodér** (*Feed Forward Neural Network Autoencoder*)
- **Konvoluční autoenkodér** (*Convolutional Neural Network Autoencoder*)
- **Transformer autoenkodér** (*Transformer Autoencoder*)

Byla použita reálná historická data z Deutsche Börse AG, která prošla *předzpracováním, extrakcí klíčových příznaků a redukcí dimenze*.

Modely byly trénovány a hodnoceny pomocí metrik **Excess Mass** a **Mass Volume**, které jsou vhodné pro neanotované datové sady. Výpočty probíhaly ve výpočetním prostředí e-INFRA CZ (MetaCentrum).

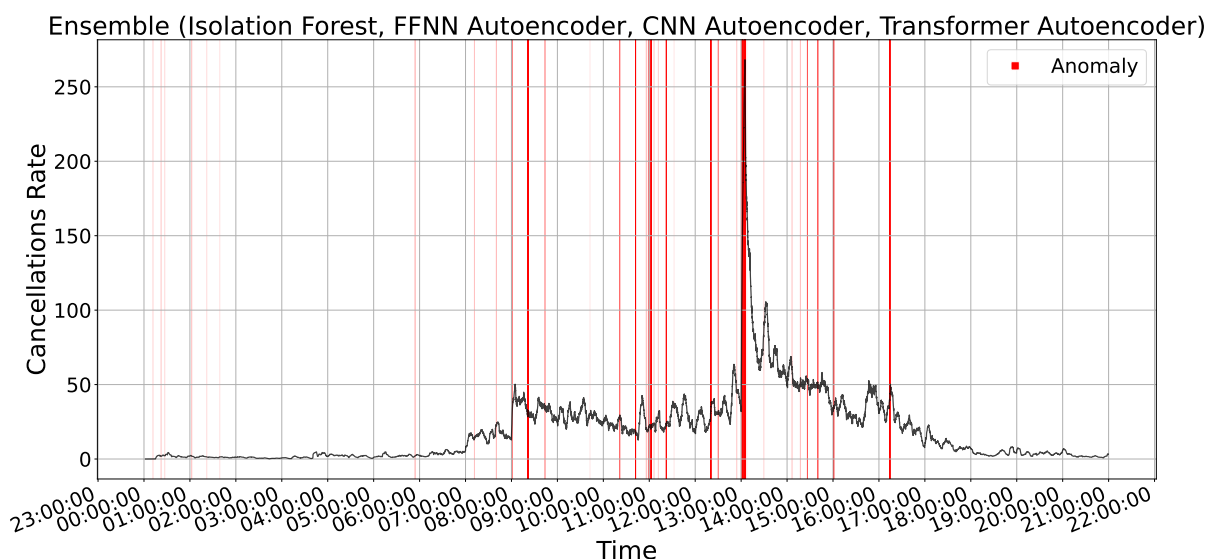
Součástí řešení je také návrh interaktivního nástroje pro vizualizaci veškerých výsledků.

5 Dosažené výsledky

Nejlepší výkonnost zaznamenaly modely *Izolační les* a *Transformer autoenkodér*. Kombinací nejlepších modelů vznikl robustní systém schopný detekovat potenciálně manipulativní chování bez nutnosti ruční anotace – viz Obrázek 1.

Navržený vizualizační nástroj má potenciál pomoci odhalit reálné případy manipulativních praktik a podporuje expertní analýzu a interpretaci detekovaných událostí.

Celý systém je plně škálovatelný a připraven pro nasazení na rozsáhlých datových sadách.



Obrázek 1: Detekované anomálie souborem modelů

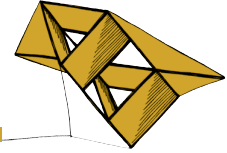
6 Závěr

Práce představuje funkční řešení pro detekci objednávek podezřelých ze *spoofingu* pomocí metod *strojového učení bez učitele*. Aplikace těchto přístupů na finanční časové řady může sloužit jako nástroj pro podporu analytiků a regulačních orgánů při odhalování tržních manipulací.

Do budoucna je možné systém rozšířit o další modely a příznaky vstupních dat. V případě dostupnosti spolehlivě anotovaných dat lze navázat také metodami *učení s učitelem*.

Literatura

- Abergel, F., Anane, M., Chakraborti, A., Jedidi, A. a Toke, I.M. (2016). *Limit Order Books*. Cambridge University Press. DOI: 10.1017/CBO9781316683040.
- Montgomery, J.D. (2016). *Spoofing, Market Manipulation, and the Limit-Order Book*. SSRN Electronic Journal. DOI: 10.2139/ssrn.2780579.



Ablations Studies in Pose-Based Sign Language Translation

Tomas Zelezny¹

1 Introduction

Sign Language Translation (SLT) task is rapidly evolving from isolated recognition approaches to a complex continuous gloss-free translation. Recent models leverage multimodal learning and large-scale language models to directly translate visual input into text. However, gloss-free SLT still faces persistent problems such as signer variability and mismatch with spoken language. This work investigates how pose-based pre-processing techniques can improve the performance of gloss-free SLT in a T5-based transformer framework.

2 Method

Data Pre-processing For our experiments, we use the YouTube-ASL (Uthus et al. (2023)) dataset for training and the How2Sign (Duarte et al. (2021)) dataset for evaluation. The YouTube-ASL dataset is uncurated and contains videos with highly variable characteristics, such as signer count, position, size, and resolution. Although How2Sign is recorded in a controlled setting, signers may still shift across videos. These inconsistencies make data pre-processing crucial. We apply our two-stage keypoint extraction pipeline, which leverages YOLOv8-nano and MediaPipe, to extract 104 keypoints consisting of the hands, body pose, and face. Then, we apply various normalization strategies to improve consistency across clips.

Model Our model is slightly modified T5 encoder-decoder transformer architecture inspired by Uthus et al. (2023). Instead of tokenized text, our model's input is a sequence of 208-dimensional vectors representing 104 pose keypoints. A custom linear embedding layer projects these vectors into the model's input space, replacing the standard text embedding layer.

3 Experiments

We conduct a series of ablation studies to evaluate how different pre-processing strategies affect the performance of our T5-based SLT model. All experiments were trained on the YouTube-ASL dataset and evaluated on the How2Sign dataset using BLEU scores as the performance metric. To address the seed variability, we repeat each experiment three times and report the best score.

Normalization We test four different normalization techniques. First two are inspired by Uthus et al. (2023) pre-processing pipeline - clip-wise ($yasl_c$) and frame-wise ($yasl_f$) normalization. Next, we propose a normalization based on the signing space ($SignSpace$). Last approach doesn't use any normalization ($none$).

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

Normalization	B-1	B-2	B-3	B-4	Method	B-1	B-2	B-3	B-4
none	13.62	3.67	1.54	0.73	none	17.47	7.19	3.79	2.17
$yasl_c$	13.00	3.90	1.59	0.66	int. ≤ 2 frames	16.91	7.35	4.06	2.43
$yasl_f$	14.67	4.78	2.19	1.13	int. ≤ 3 frames	17.16	7.40	4.01	2.33
<i>SignSpace</i>	17.47	7.19	3.79	2.17	light aug.	15.76	6.23	3.12	1.71
					medium aug.	17.27	7.51	4.12	2.46
					heavy aug.	16.58	7.10	3.85	2.29

Table 1: Comparison of four different types of normalization techniques, interpolation and augmentations. Performance is measured by BLEU scores on the How2Sign dataset.

Interpolation Next, we experiment with linear interpolation of missing keypoints instead of using default value -10. We experiment with a total of 3 different settings: interpolate all gaps with size 2 or smaller, with gaps 3 or smaller, or don’t use interpolation at all.

Augmentation In our final set of experiments, we investigate various types and intensities of augmentation. Our preliminary experiments revealed that applying to the input image, rotating pose in the elbow joints, applying gaussian noise have beneficial effect on the performance. Based on these findings, we designed three augmentation protocols with increasing strength, combining these three techniques.

4 Conclusion

This work systematically investigates the impact of pose-based pre-processing techniques on the gloss-free Sign Language Translation using a T5-based transformer model. Through extensive ablation studies, we show that effective handling of the input plays an important role in translation quality. Our proposed SignSpace normalization achieves the best performance, while interpolation proves to be essential for addressing missing keypoints. Additionally, data augmentation further boosts model robustness and accuracy. Our code is public at our GitHub¹.

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

- Uthus, D., Tanzer, G., Georg, M. (2023). Youtube-asl: A large-scale, open-domain american sign language-english parallel corpus. *Advances in Neural Information Processing Systems*, 36, 29029-29047.
- Duarte, A., Palaskar, S., Ventura, L., Ghadiyaram, D., DeHaan, K., Metze, F., ... & Giro-i-Nieto, X. (2021). How2sign: a large-scale multimodal dataset for continuous american sign language. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* pp. 2735-2744.

¹https://github.com/zeleznyt/T5_for_SLT/tree/cb5fa58907b34365316f017ce6fe13d0116c829b