

Pokročilé získávání datasetů pro přesnou klasifikaci stavebního a demoličního odpadu pomocí strojového učení

**Tomáš Zbiral¹, doc. Ing. Václav Nežerka, Ph.D.¹, Ing. Jan Trejbal, Ph.D.¹, Matěj Hužvár²,
doc. Ing. Stanislav Vitek, Ph.D.²**

¹České vysoké učení technické v Praze, Fakulta stavební, Thákurova 2077/7, 160 00 Praha 6,
E-mail: jan.trejbal@fsv.cvut.cz

²České vysoké učení technické v Praze, Fakulta elektrotechnická, Technická 1902/2,
160 00 Praha 6

Summary

The dependable sorting of distinct materials from construction debris stands as a foundational element within the circular economy in construction industry. This paper introduces a technology that employs RGB cameras, mass, and acoustic sensors for material classification. By amalgamating these sensors and applying attribute extraction techniques such as shape indices, texture entropy, and intensity gradients, a comprehensive set of parameters is generated. These parameters facilitate accurate classification through the utilization of artificial intelligence.

Keywords: Construction and Demolition Waste, Waste Sorting, RGB Cameras, Mass Sensor, Acoustic sensor, Material Classification, Artificial Intelligence

Úvod

Stavební a demoliční odpad (SDO) globálně vytváří významnou zátěž pro ekonomiku a životní prostředí. Zpravidla se jedná o různorodou směs stavebních materiálů s potenciálem dalšího využití. Jeho efektivní recyklace je ale podmíněna velmi přesným tříděním. Doposud rozšířené metody ovšem v tomto ohledu často selhávají. Jsou technologicky a časově náročné, pracné, nespolehlivé a v řadě případů i nebezpečné pro obsluhující personál [1]. Zejména v posledních třech dekádách tak roste zájem o vývoj automatických třídících linek, které zmiňované nedostatky v co největší míře potlačují.

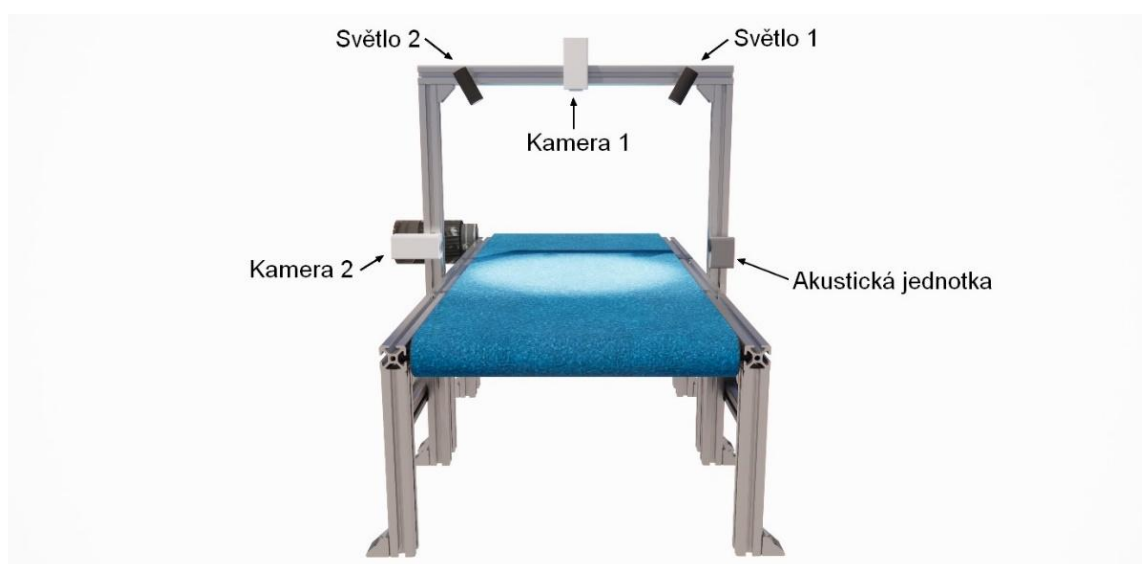
Velmi slibných výsledků dosahují třídící linky, které pro identifikaci a třídění jednotlivých materiálů využívají RGB kamery a algoritmy strojového učení (machine learning). Pracují na bázi hlubokého učení (deep learning) a konvolučních neuronových sítí (convolutional neural networks). Pilotní pokusy s jejich uplatněním při třídění SDO během posledních třech let prezentovaly především čínské výzkumné týmy. Ukázalo se, že tyto metody jsou rychlé a velmi účinné, ovšem pouze v případě materiálů s jasně definovatelnou texturou povrchu [2, 3]. Vizualní data často nejsou dostatečně určující, čímž dochází k jejich splynutí a chybné klasifikaci. Další nepřesnosti do procesu třídění vnáší různá znečištění, prašnost a ostatní artefakty potlačující charakteristický povrch konkrétního materiálu. Tomuto problému se lze v praxi vyhnout jen velmi obtížně. Nabízí se ale aplikace multisenzorového sběru dat, které vizualní informace doplní o další měřitelné veličiny.

Pro překonání výše zmiňovaných limitů byla navržena pilotní verze třídící linky, která kromě RGB kamer synergicky využívá akustické senzory a zařízení pro měření hmotnosti. Sensorový modul je integrován s pásovým dopravníkem pro transport drtě SDO. Tato sestava usiluje o vytvoření velmi komplexního souboru vstupních dat, která jsou zpracovávána modelem strojového učení. Každému fragmentu SDO je přiřazena množina informací, která pokrývá atributy vizualní (barva, textura, entropie), akustické (odezva na ultrazvukový impuls) a hmotnostní (společně s hodnocením vizualního charakteru dochází k výpočtu objemové hmotnosti). Celý modul lze navíc rozšířit o další typy senzorů, např. hyperspektrálním snímkováním. Hlavním cílem našeho úsilí je vyvinout robustní, vysoce přesný

klasifikační systém, který je schopný rozlišovat mezi různými typy SDO. V případě úspěchu lze očekávat zvýšení efektivity třídících procesů se značným potenciálem uplatnění v praxi.

Vývoj multimodálního systému sběru dat

Ústřední součástí třídícího systému je pásový dopravník, který zajišťuje kontinuální tok fragmentů SDO, viz Obrázek 1. Dopravní pás zajišťuje konzistentní prezentaci materiálu k sensorovému modulu a udržuje pravidelnou vzdálenost a orientaci mezi jednotlivými fragmenty, což je zásadní předpoklad pro získání spolehlivých a reprodukovatelných údajů. Nad dopravníkovým pásem je umístěn pomocný rám pro osazení příslušenství. Na něm je upevněna dvojice kamer RGB, která snímá fragmenty SDO ve dvou rovinách. Rám je dále vybaven nastavitelným zdrojem světla pro zajištění konzistentních světelných podmínek při různých snímacích relacích. Kromě vizuálních informací systém integruje akustický senzor a zařízení pro měření hmotnosti, čímž znatelně rozšiřuje spektrum shromažďovaných dat. Zařízení pro měření hmotnosti pracuje na základě okamžité síly působící na válečky, které vedou dopravní pás. Poskytuje tak vypovídající údaje o hmotnostních charakteristikách různých fragmentů SDO.



Obrázek 1: Prototyp vyvinutého multimodálního akvizičního systému.

Akustická jednotka materiál klasifikuje na základě detekce ultrazvukových vln odražených od povrchu zkoušeného materiálu. Pro měření jsou použity ultrazvukové snímače HC-SR04, přičemž jejich umístění maximalizuje vyzařovací charakteristiky směrem k cílovému vzorku. Ovládány jsou pomocí mikrokontroléru Arduino UNO připojeného k počítači. Kromě toho je na pomocný rám připevněn měřicí mikrofón Brüel & Kjær 1/8", který snímá signál odražený od povrchu vzorku. Získaný signál se dále zpracovává pomocí měřicího zesilovače Brüel & Kjær Nexus s nastavením převodu 100 mV/Pa. Následně se zesílený signál zobrazí na digitálním osciloskopu Rigol MSO5074.

Vzorky použité pro trénink identifikace

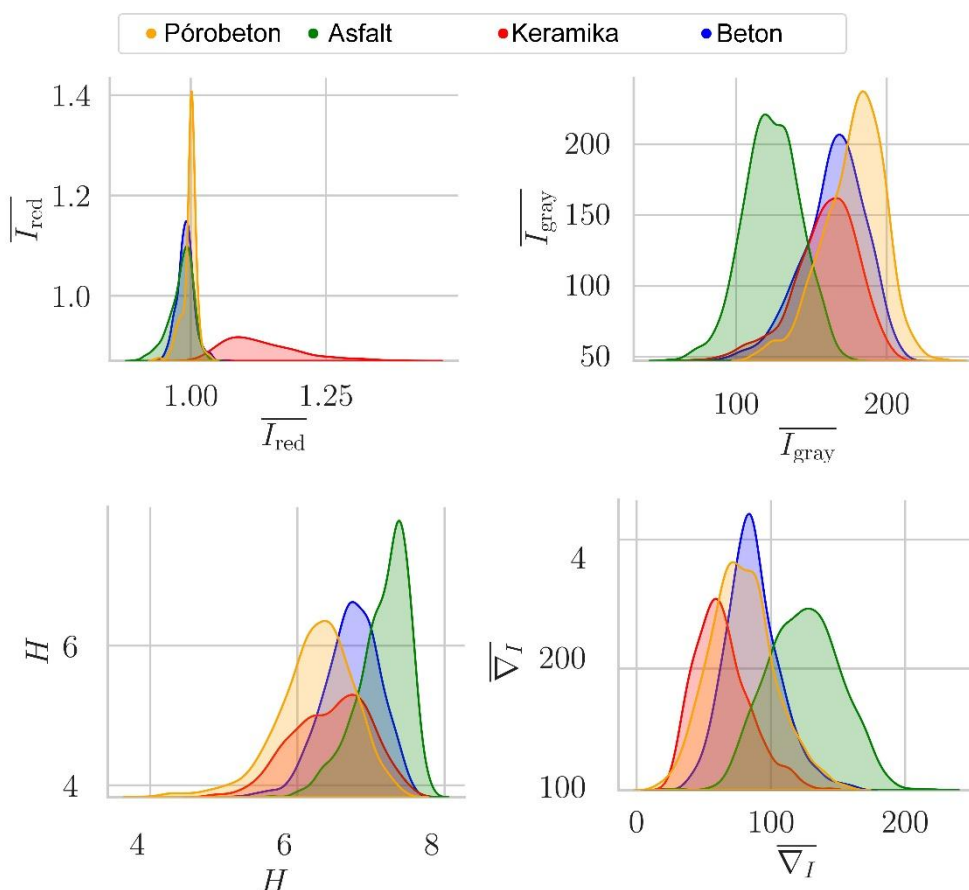
Pro účely vývoje třídícího systému byly použity 4 reprezentativní typy SDO: pórobeton (AAC), asfaltový beton (Asphalt), keramické cihly a střešní tašky (Ceramics) a cementový beton (Concrete). Každý z těchto materiálů byl zastoupen fragmenty o rozměrech 50 až 200 mm, viz Obrázek 2.

Obrázek 3 zobrazuje korelaci rysů extrahovaných z RGB dat dle metody vyvinuté na Fakultě stavební ČVUT, viz [4]. Veličina $\overline{I_{red}}$ značí střední intenzitu červené barvy v poměru k průměrnému jasům snímku.

Zjednodušeně řečeno určuje množství červené barvy zastoupené na snímku. Podobného, avšak obecnějšího významu dosahuje veličina $\overline{I_{gray}}$, která definuje celkový jas snímku. Aby s sebou nesla vypovídající hodnotu, je nutné snímky pořizovat za konstantních světelných podmínek a při konstantním nastavení expozice. Podrobnější informace poskytuje veličina H , která se značí jako Shannonova entropie (Shannon's Entropy). Tato charakteristika přehlíží rozložení barevných kanálů a jejich intenzit, zatímco se soustředí na popis textury snímaného vzorku. Čím vyšších hodnot dosahuje, tím vyšší je náhodnost zaznamenaného vzoru. Poslední analyzovanou veličinou pomocí RGB kamer je tzv. gradient střední intenzity $\overline{\nabla I}$ (Mean Intensity Gradient). Jedná se o indikátor stochastického vzoru, který definuje hrubost povrchu.



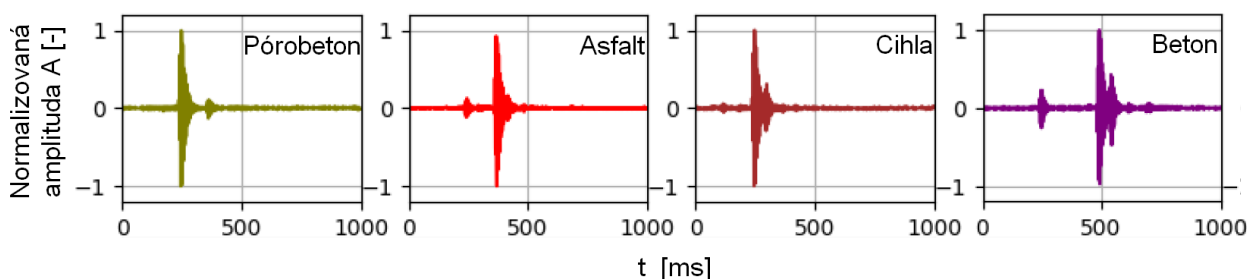
Obrázek 2: reprezentativní digitální snímky analyzovaných fragmentů SDO s vyznačenou oblastí zájmu, zleva: pórobeton, asfalt, cihla a beton.



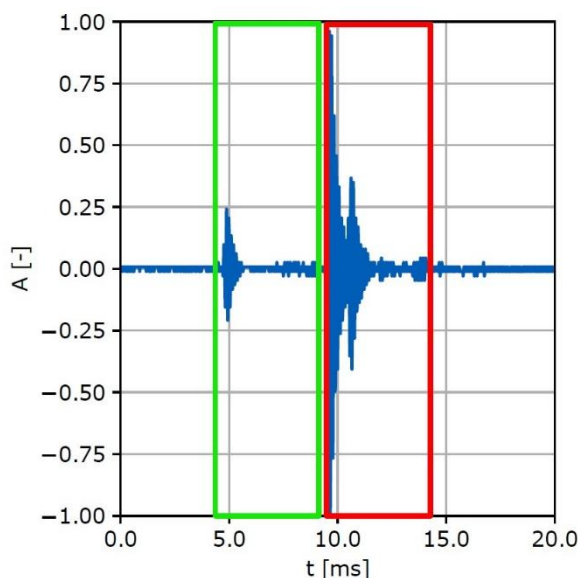
Obrázek 3: Korelace rysů extrahovaných z obrázků RGB [4].

Pokud jde o akustickou analýzu SDO, naměřené signály odražených ultrazvukových vln vykazují krátký úzkopásmový signál na frekvenci přibližně 40 kHz. Na začátku byly signály normalizovány

odečtením jejich střední hodnoty. Ačkoli tento postup normalizace eliminuje potenciální informace týkající se akustické absorpce materiálu, vzhledem k zamýšlené robustnosti klasifikační metody v nekalibrovaných polních podmínkách se považuje za přijatelný. Stojí za zmínku, že všechny signály vykazují výrazný amplitudový vrchol, po němž následuje menší vrchol, kterému často předchází slabší aktivita. Pro usnadnění komplexnější analýzy byla zvolena dvě časová okna. Data odezvy shrnuje Obrázek 4 níže. Obrázek 5 přibližuje detail zaznamenané odezvy. Předmětnou oblast dělí do dvou oken, ve kterých je znatelná unikátní odezva pro každý ze zkoumaných materiálů. Zelené okno zahrnuje oblast před dominantním vrcholem. Červené okno se soustředí na popis tvaru dominantního píku a následujících shluků.



Obrázek 4: Variabilita odražených signálů naměřených pro různé materiály.



Obrázek 5: Okna pro extrakci rysů z načteného ultrazvukového signálu.

Fúze dat z různých senzorů

Efektivní využití multimodálních dat v klasifikační úloze vyžaduje nejen jejich sběr z různých senzorů, ale také následné logické slučování. Samotný proces fúze kombinuje informace z jednotlivých senzorů za využití unikátních parametrů typických pro každý analyzovaný vzorek. Tímto přístupem jsou eliminovány nepřesnosti a informační mezery, které jsou dány limity jednotlivých senzorových technik. V kontextu klasifikace SDO to představuje integraci informací získaných ze snímků RGB, měření akustické odezvy a měření hmotnosti za účelem přesnějšího rozlišení různých typů fragmentů SDO.

Prvním krokem tohoto procesu je nezávislá extrakce příznaků z dat každého senzoru. V případě vizuálních dat je použita charakterizační metodika, které se věnuje samostatná publikace [4]. Založena je na extrakci různých texturních rysů z obrazového snímku s vysokým rozlišením. Oproti tomu pro

akustická data jsou požitý standardní techniky zpracování signálu pro extrakci spektrálních a časových rysů, jak uvádí předchozí kapitola.

Po extrakci souborů funkcí ze všech použitých senzorů následuje jejich fúze. Strategii slučování dat lze dělit do třech následujících úrovní: včasné slučování (slučování na úrovni dat), střední slučování (slučování na úrovni prvků) a pozdní slučování (slučování na úrovni rozhodnutí).

Při včasné fúzi se surová data z jednotlivých senzorů kombinují před jakýmkoli zpracováním nebo extrakcí funkcí. Vzhledem k rozdílné povaze vizuálních a akustických dat však v prezentovaném případě není časná fúze proveditelná.

Při meziproductové fúzi se sady příznaků z jednotlivých senzorů spojí do jediného vysoko-dimenzionálního vektoru, který se poté vloží do klasifikačního modelu. Tento přístup umožňuje klasifikátoru rozhodovat na základě všech dostupných informací, ale také zvyšuje dimenzionalitu vstupních dat, což může vést k problémům spojeným s takzvaným "prokletím dimenzionality" (Curse of Dimensionality).

Při pozdní fúzi je sada funkcí každého senzoru nezávisle vložena do samostatných klasifikátorů a konečné rozhodnutí je učiněno na základě kombinace těchto samostatných rozhodnutí. Tento přístup využívá silné stránky každého senzoru při rozpoznávání určitých typů fragmentů SDO, ale při rozhodování se do značné míry spoléhá na robustní strategii.

V této studii byl použit přístup meziproductové fúze. Oproti ostatním metodám vyniká ve schopnosti využít komplementární povahu vizuálních a akustických funkcí. Sloučený vektor vysoko-dimenzionálních příznaků je předán klasifikátorům strojového učení. Informace jsou následně podrobeny algoritmu rozhodovacích stromů (Tree Decision) s gradientním posilováním (Gradient Boosting, GB) a vícevrstevnými perceptrony (Multilayer Perceptron, MLP).

Trénování modelů strojového učení pro klasifikaci SDO

Prvním krokem v procesu trénování je rozdělení celé sady dat na dvě podmnožiny: tréninkovou a ověřovací sadu. Tréninková množina slouží k učení modelu, zatímco validační množina se používá pro testování přesnosti modelu na neznámých datech a k ladění hyperparametrů. Standardní praxí je použít přibližně 70-80 % datové sady pro trénování a zbývajících 20-30 % pro validaci.

Proces trénování zahrnuje úpravu parametrů modelu tak, aby jeho předpovědi na základě tréninkových dat byly co nejpřesnější. Toho je dosaženo definováním ztrátové funkce, která kvantifikuje rozdíl mezi předpověďmi modelu a skutečností. Parametry modelu se pak iterativně aktualizují pomocí optimalizačních algoritmů, mezi které patří například gradientní sestup. Postupně se tak minimalizuje ztráta.

V případě modelů GB i MLP zahrnuje proces učení několik kol iteračního učení. U prvně zmíněného se trénuje řada rozhodovacích stromů, přičemž každý následující krok se učí opravit chybu toho předchozího. Druhý model využívá odlišný přístup. Učí se mapovat vstup (rysy) na výstup (štítky) pomocí několika vrstev umělých neuronů (perceptorů), přičemž v průběhu procesu učení jsou aktualizovány jejich váhy.

Po dokončení trénování se výkonnost modelu vyhodnotí na validační množině. Tento krok je klíčový pro zajištění schopnosti modelu dobře zobecnit neznámá data a zabránit nadměrnému přizpůsobení. Může se totiž stát, že model funguje spolehlivě na tréninkových datech, ale nová data hodnotí chybně.

Jakmile jsou modely dobře natrénovány a prokáží uspokojivou výkonnost na validačním souboru, mohou být nasazeny pro klasifikaci fragmentů SDO v reálném čase. Využití těchto algoritmů lze předpokládat v třídících SDO a recyklačních závodech.

Přechod na průmyslové třídění SDO

Veškeré popisované experimenty se odehrávaly v laboratorním prostředí. Přejít na průmyslové měřítko si bude v první řadě žádat kapacitní a rychlý dopravníkový systém pro odbavení velkého množství SDO. Nezbytnou podmínkou je též instalace technologie pro prostorovou separaci jednotlivých fragmentů SDO tak, aby mohl být každý vzorek jasně ohraničen. Dalším předpokladem pro průmyslovou implementaci je osazení většího množství skupin čidel, a to na více místech pásového dopravníku. Nezbytnou součástí se musí stát dostatečně výkonná výpočetní sestava, které bude schopna v reálném čase vyhodnocovat veškeré toky dat.

Před průmyslovou aplikací bude též nutné zpracovat a analyzovat podstatně větší množství dat pro trénování modelů strojového učení. Proces učení ale může probíhat již za skutečného provozu. Obecně platí, že čím větší množství dat model hodnotí, tím vyšší přesnosti dosahuje. Prakticky aplikovaný model musí být odolný vůči odchylkám v typech SDO, světelných podmínkách, kalibraci senzorů a mnoha dalším stochastickým proměnným.

Závěr

Prezentovaná studie popisuje pokročilý přístup k rozpoznávání a třídění SDO. Využívá fúze dat získaných z RGB obrazových, ultrazvukových a hmotnostních senzorů. Takto získaná data jsou následně hodnocena pomocí modelů strojového učení. Synergické využití multisenzorové techniky do značné míry eliminuje omyly, kterých by se jednotlivá čidla dopouštěla při samostatné práci. Perspektiva použití těchto metodik v průmyslovém měřítku představuje zajímavé možnosti pro oblast nakládání nejen s SDO a jejich recyklace. Nutno dodat, že přechod do průmyslového měřítko bude vyžadovat těsnou spolupráci se zástupci z průmyslového sektoru, vládními agenturami a dalšími zúčastněnými stranami.

Poděkování

Tato práce byla podpořena rámcovým programem Evropské unie Horizon Europe (výzva HORIZON-CL4-2021-TWIN-TRANSITION-01-11) na základě grantové dohody č. ~101058580, projektem RECONMATIC (Automatizovaná řešení pro udržitelné a oběhové hospodářství stavebních a demoličních odpadů), Technologickou agenturou České republiky, grantovou dohodou č. S03010302 (Vývoj účinných nástrojů pro minimalizaci produkce stavebních a demoličních odpadů, jejich sledování a opětovné využití) a ČVUT v Praze, grantová dohoda č. SGS23/004/OHK1/1T/11 (Vývoj a aplikace digitalizovaných a automatizovaných nástrojů pro víceúčelové aplikace ve stavebnictví).

Reference

- [1] Hoong, J.D.L.H., Lux, J., Mahieux, P.Y., Turcry, P., & Aït-Mokhtar, A. 2020. "Určení složení recyklovaných kameniv pomocí analýzy obrazu založené na hlubokém učení". *Automatizace ve stavebnictví* 116: 103204. <https://doi.org/10.1016/j.autcon.2020.103204>.
- [2] Ku, Y., Yang, J., Fang, H., Xiao, W., & Zhuang, J. 2020. "Hluboké učení detekce uchopení pro robota používaného při třídění stavebního a demoličního odpadu". *Journal of Material Cycles and Waste Management* 23: 84-95. <https://doi.org/10.1007/s10163-020-01098-z>.
- [3] Lin, K., Zhou, T., Gao, X., Li, Z., Duan, H., Wu, H., Lu, G., & Zhao, Y. 2022. "Deep Convolutional Neural Networks for Construction and Demolition Waste Classification" (Hluboké konvoluční neuronové sítě pro klasifikaci stavebního a demoličního odpadu): VGGNet Structures, Cyclical Learning Rate, and Knowledge Transfer" (Struktury VGGNet, cyklická míra učení a přenos znalostí). *Journal of Environmental Management* 318: 115501. <https://doi.org/10.1016/j.jenvman.2022.115501>.
- [4] Nežerka, V., Zbírál, T., & Trejbal, J. 2023. "Strojově učící se klasifikace fragmentů stavebního a demoličního odpadu pomocí počítačového vidění: Convolution versus Extraction of Selected Features". Předáno do časopisu *Automatizace ve stavebnictví*.