

Česká zemědělská univerzita v Praze

Fakulta agrobiologie, potravinových a přírodních zdrojů

Katedra agroekologie a rostlinné produkce



**Fakulta agrobiologie,
potravinových a přírodních zdrojů**

**Využití metod strojového učení pro detekci plevelů v
plodinách**

Diplomová práce

Mgr. et Bc. Adam Hruška

Pěstování rostlin

Vedoucí práce Ing. Pavel Hamouz, Ph.D.

© 2022 ČZU v Praze

Čestné prohlášení

Prohlašuji, že svou diplomovou práci "Využití metod strojového učení pro detekci plevelů v plodinách" jsem vypracoval(a) samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu literatury na konci práce. Jako autor(ka) uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 14.4.2022

Poděkování

Rád bych touto cestou poděkoval Ing. Pavlu Hamouzovi Ph.D. za jeho velkou pomoc s tvorbou analýzy od jejích prvních počátků až k úplnému závěru. V druhé řadě bych chtěl poděkovat Ing. Jakobovi Lvovi Ph.D. za konzultaci a podporu v technických a praktických aspektech práce. V neposlední řadě bych velmi rád poděkoval celé své rodině a partnerce Pětě za neustálou psychickou a motivační podporu při studiu a tvorbě závěrečné práce.

Využití metod strojového učení pro detekci plevelů v plodinách

Souhrn

Primárním cílem práce bylo ověření možností využití morfologických znaků plevelných rostlin pro jejich detekci na bázi strojového učení. Pro dosažení cíle bylo nutné nasbírat dostatečný objem relevantních dat, které následně sloužily jako báze pro trénink a evaluaci modelu strojového vidění. Příprava vhodné platformy pro sběr dat, společně se systematickou přípravou získaných snímků pro trénink modelu, tvořila dílčí cíle práce. Jednotlivé procesy analýzy byly realizovány na základě empirických poznatků z obdobně směřovaných výzkumů a kontextu lokální agrární sféry. Druhotnými cíli práce bylo v první řadě ověření výstupů vlastního modelu tréninkem v současnosti prakticky používaného modelu YOLOv3, v druhé se jednalo o empirické ověření problematiky praktického využívání technologie v ekonomické a technologické rovině.

Na základě empirických poznatků bylo pro účely analýzy vybráno rozhraní infračerveného pásma snímků, které nabízelo efektivnější oddělení snímaných objektů od jejich pozadí. Snímky samotné zachycovaly 18 druhů plevelných rostlin, které lze v kontextu polního hospodářství v České republice považovat za signifikantní pro ovlivnění výnosu plodin, zejména okopanin. Jednotlivé plevelné rostliny byly na snímcích v rámci přípravy ručně anotovány a výsledný dataset prošel cílenými úpravami zodpovědnými za efektivní zvětšení objemu dat. Před počátkem tréninku byla data náhodně rozdělena na tréninková, evaluační a testovací v poměru 70:10:20. Samotný trénink obou modelů následoval jednotnou logiku vstupu připravených dat a následného statistického ověření výsledků.

Výsledky obou modelů překročily hranici 90 % přesnosti, kdy vlastní model dosahoval na testovacích datech 94,5 % úspěšnosti detekce při rychlosti zpracování 28 ms na snímek. Model YOLOv3 ve srovnání dosáhl hodnot 97,92 % úspěšnosti při rychlosti zpracování 22,8 ms na snímek. Ačkoliv se jednalo o znatelné rozdíly ve výsledcích, tak aplikace modelu YOLOv3 prakticky potvrdila fungování vlastního modelu vzhledem k jeho experimentálnímu formátu. Empirické posouzení aplikovatelnosti modelů do prostředí praxe na jedné straně potvrdilo reálnost implementace, ovšem na straně druhé vyzdvihlo i řadu slabých stránek a hrozeb pro fungování technologie.

Klíčová slova: detekce plevelů, RGB snímky, strojové učení, analýza obrazu, neuronové sítě

Using of machine learning methods for weed detection in crops

Summary

The primary goal of the analysis consists of the use of morphological features of weeds for their detection through machine learning. For successful completion of the analysis, the computer vision model required immense amounts of relevant data for training and evaluation. The preparation of a data-gathering platform as well as the systematic preparation of the obtained images were the main sectional goals of the thesis. The processes of the analysis were conducted on the basis of the empirical knowledge from similar research and the context of local agricultural industry. Secondary goals of the analysis included the comparison of the model through training of the state-of-the-art model YOLOv3 on the same data. And the empirical evaluation of the potential for the practical implementation of the technology in the context of economic and technological elements.

Empirical data led to the choice of the near infrared spectrum of the images, thanks to its more effective elimination of the background from the targeted objects. The data included 18 species of weeds, which in the context of the agricultural industry of the Czech Republic pose a significant threat to crop yields, especially in case of root crops. Each image was manually annotated, and the final dataset went through selected adjustments effectively doubling the amount of data. The dataset was randomly split into three parts for training, evaluation and testing with the ratio of 70:10:20 prior to the training phase. The training itself followed identical approaches for both models with the training and subsequent statistical evaluation of the results.

Both models achieved more than 90 % accuracy. The testing of the original model resulted in 95.5 % accuracy in detection and identification with the processing speed of 28 ms per image. The YOLOv3 achieved 97.92 % accuracy and speed of 22.8 ms per image. Despite the differing results, the application of YOLOv3 validated the original model because of its experimental format. The empirical analysis resulted in the confirmation of the potential practical implementation of the technology, but at the same time highlighted the weaknesses and potential threats to the operation of the computer vision in agriculture.

Keywords: Weed detection, RGB pictures, Machine learning, visual analysis, neural networks

Obsah

| | |
|---|-----------|
| 1 Úvod | 7 |
| 2 Vědecká hypotéza a cíle práce | 9 |
| 3 Literární rešerše..... | 10 |
| 3.1 Báze strojového učení | 10 |
| 3.1.1 Metody v rámci strojového učení | 12 |
| 3.1.2 Neuronové sítě YOLO | 15 |
| 3.2 Koncepční prvky | 17 |
| 3.2.1 Obecně koncepční články | 17 |
| 3.2.2 Publikace koncipující vlastní projektové řešení | 18 |
| 3.3 Empirický potenciál pro praktickou implementaci..... | 22 |
| 3.3.1 Ekonomický efekt..... | 22 |
| 3.3.2 Odolnost k nekontrolovaném prostředí | 24 |
| 3.4 Směřování výzkumného rámce..... | 26 |
| 4 Metodika | 27 |
| 4.1 Sběr dat..... | 27 |
| 4.2 Příprava a anotace snímků..... | 28 |
| 4.3 Tvorba modelů a jejich trénink | 30 |
| 4.3.1 Vlastní model..... | 30 |
| 4.3.2 YOLOv3 | 33 |
| 4.3.3 Statistické zhodnocení úspěšnosti..... | 33 |
| 5 Výsledky..... | 34 |
| 5.1 Výsledky vlastního modelu..... | 34 |
| 5.2 Výsledky modelu YOLOv3 | 35 |
| 6 Diskuze | 37 |
| 6.1 Komparace modelů | 37 |
| 6.2 Srovnání s obdobnými modely..... | 39 |
| 7 Závěr | 41 |
| 8 Literatura..... | 43 |

1 Úvod

Soudobá agrární sféra podléhá v kontextu globálního dění řadě silných vlivových proudů. Mezi nejvýznamnější patří ekonomika potravinové produkce, enviromentální zodpovědnost a potravinová bezpečnost. Přestože individuální vliv jednotlivých faktorů na zemědělství v postupu času formuje jeho celkovou koncepci, absolutní převaha individuálního směru není v zásadě realistická vzhledem k jejich zaměření (Guerin & Guerin 1994). Všechny hlavní proudy ovšem nacházejí v zásadních oblastech průsečíky svých zájmů, které následně vedou k zásadním změnám v legislativním i praktickém přístupu k potravinové produkci. Z historického hlediska lze připomenout poválečné časy nedostatku, které silným ekonomickým a bezpečnostním tlakem zásadně pomohly v rozvoji chemické ochrany plodin, která vyústila v bezprecedentní nárůst výnosů (Lahsen 2022). Kvantitativně se implementace chemických přípravků do oboru projevila i následným růstem objemů produkce, kdy výstupy narostly o neuvěřitelných 145 % od počátku šedesátých let (Pretty 2008).

V současnosti je ovšem navzdory dřívějšímu podílu na řešení nesčetných hrozeb chemická ochrana sama problémem řešeným v rámci jednoho z výše zmíněných průsečíků. I přes jeho elementární roli poskytovatele potravin je dnes agrární sektor posouván do dalších, mnohdy značně inovativních pozic, které pramení z nově formulovaných potřeb společnosti. Vedle základní potřeby výživy dochází k zásadnímu tlaku na dlouhodobou udržitelnost produkce, a to jak v enviromentálním a společenském, tak v ekonomickém hledisku (Pretty 2008). Mezi její manifestace patří i snaha o zachování diverzity prostředí, která dlouhodobě posiluje stabilitu ekosystémů. Na jedné straně tedy existuje požadavek na omezení chemických zásahů v relevanci k minimalizaci škody na necílových organismech, ovšem na straně druhé zůstává nutnost dostatečného zásobování potravinami (Schaller 1993).

Samotné zásobování navíc čelí problematice rostoucí populace a jejích markantně rostoucích potřeb (Pretty 2008), střetávající se s postupným úbytkem neobnovitelných zdrojů nezbytných pro fungování současné potravinové produkce (Pfeiffer 2006). Dle většiny odhadů populační růst dosáhne svého pravděpodobného vrcholu na sklonku 21. století, kdy by objem lidské populace měl překonat hranici deset miliard (Lutz et al. 2001). Navzdory relativní stabilitě trendu a jeho případným výkyvům není v současnosti potravinová produkce schopna saturovat potřeby narůstající s populačním příbytkem. Odhady potřebného nárůstu produkce se pohybují v mezích 70 – 110 % současných výstupů, přičemž variabilita podléhá jak nejistotě lineárního růstu, tak změnám potřeb soudobé společnosti (Pretty 2008).

V rovině klesající dostupnosti a množství neobnovitelných zdrojů se jako jeden z primárních problémů jeví absolutní závislost oboru na ropných produktech a zemním plynu (Pfeiffer 2006). Ačkoliv většina projekcí prezentuje relativně velký časový rámec do vyčerpání existujících nalezišť, není možné na 30 - 40leté období pohlížet jako na limit pro implementaci alternativ. Samotná závislost na komoditách sekundárně oslabuje stabilitu systému i skrze produkci skleníkových plynů působící na vývoj klimatických změn. Celkově je dle odhadů na jeden hektar produkce nutné využít zhruba 1000 litrů nafty v kontextu energetické náročnosti agrotechniky, dopravy i výroby přípravků na ochranu rostlin a hnojiv (Pfeiffer 2006). Vedle ropy se jako zásadní hrozba jeví i zemní plyn, jehož význam pro produkci chemických prostředků je v současné době prakticky nenahraditelný. V případě pokračování aktuálních trendů lze v relativně blízké budoucnosti očekávat nedostatky v rovině pohonných hmot, energetických surovin a minerálních hnojiv (Pfeiffer 2006).

Vzhledem k souběžnosti jednotlivých problémů se jako potenciální řešení nabízí zefektivnění využití dostupných chemických prostředků v rámci precizního zemědělství (Gebbers & Adamchuk 2010). Rámcová koncepce počítá s využitím přesných senzorů, jakožto mediátorů lokace a objemu aplikované dávky chemického prostředku, který by díky nově získaným hranicím dosahoval komparativně obdobných výsledků, jako při plošném užití. Ač současné zpřesnění aplikací pomocí navigačních systémů nabízí znatelné zlepšení, dosažení společensky vytyčených cílů vyžaduje mnohonásobně úspěšnější přístup (Schaller 1993). Poslední vývoj v akademické sféře poukazuje na potenciál vizuálně řízených platforem pro dosažení prezentovaných cílů v širokém spektru podmínek.

Implementace technologie strojového vidění (computer vision) do praxe agrární sféry je i v současnosti stále limitovaná, ovšem potenciál pro rozvoj strojového učení v oboru zůstává zásadním pro naplnění cílů udržitelnosti. V jádru koncepce leží princip zpřesnění lokace a objemu dávky chemického prostředku v relevanci k cílové plevelné rostlině, která je s relativně vysokou přesností detekována integrovaným systémem (Liakos et al. 2018). Strojové učení nabízí pro koncipovanou technologii sadu nástrojů umožňující jak vysokou přesnost detekce, tak vysokou rychlost zpracování získaných dat (Sharma et al. 2020). V praxi tedy technické řešení může dosáhnout až zásadního snížení potřebných chemikálií pro ošetření porostu plodiny bez zásadnějšího rizika pro necílové organismy.

2 Vědecká hypotéza a cíle práce

Hlavním cílem práce je identifikace vhodných rozlišovacích znaků plevelů a tvorba analytického nástroje schopného detekce plevelných rostlin v polních podmínkách na základě vizuálních dat. Základním dílčím cílem projektu je tvorba platformy schopné zajistit snímkování plevelných rostlin v dostatečné kvalitě i kvantitě pro další zpracování. Druhým dílčím cílem analýzy je úspěšná implementace senzoru sbírajícího vizuální data na výše zmíněnou platformu. Třetím dílčím cílem je sběr samotných dat a jejich následná příprava a anotování pro úspěšné trénování finálního algoritmu. Finálním dílčím cílem, skrze který má práce dosáhnout svého hlavního cíle, je tvorba a následný trénink algoritmu strojového učení schopného dosažení podmínek vytýčených v hlavním cíli práce.

H₁: Při využití neuronových sítí je možné detekovat plevelné a kulturní rostliny s jistotou odlišení přesahující 85 %.

H₂: Algoritmus strojového učení je schopen vzájemně rozlišit vybrané plevelné druhy s přesností přesahující 85 %.

V širším kontextu předpokládaných cílů a hypotéz práce jsou pro výslednou analýzu přiřazeny dva sekundární cíle. V první řadě se jednalo o srovnávací ověření výstupního algoritmu skrze trénink současného prakticky používaného modelu strojového vidění. Dílčími cíli v procesu jsou výběr odpovídajícího modelu, jeho implementace a výsledné srovnání výstupů obou modelů při užití identických dat. Druhým ze sekundárních cílů je rámcový empirický rozbor problematiky praktického využívání technologie. Cíl samotný se rozděluje do dvojice dílčích rovin, kdy první z nich nahlíží na problematiku z ekonomického hlediska a jeho komplexních vztahů. Druhou dílčí rovinou cíle je empiricko-teoretický rámec fungování samotné technologie v praktických polních podmínkách, které jsou vzhledem k fungování výstupního nástroje zcela zásadní.

3 Literární rešerše

Jádro akademické literatury užitá k obohacení práce obsahuje poměrně variabilní kombinaci textů, které ovšem spojuje základní paradigma rozvoje precizního zemědělství s užitím technik strojového učení. Při bližší specifikaci je možné relevantní literaturu podoboru detekce plevelů rozdělit do dvou primárních skupin, tedy obecně koncepční texty zabývající se rozvojem zmíněné technologie a projektově implementační literaturu poskytující metodickou základnu a kontext jednotlivých prvků výzkumu. Úvod tvoří kapitola nabízející náhled do statisticko-technologické báze technologie strojového učení a jeho teoretických základů, včetně jednotlivých metodických přístupů k disciplíně zakončených vybraným soudobým modelem určeným pro srovnání výsledků. Druhou část kapitoly tvoří výše zmíněné koncepční články zabývající se implementací technologie do agrární sféry. Jednotlivé texty jsou řazené v chronologickém a technologickém postupu, kdy je teoretická báze zmíněna na úvodu podkapitol a výsledek soudobého poznání na jejich závěru. Finální podkapitola obsahuje empirický pohled na prvky ovlivňující potenciál implementace technologie doplněné o souhrn výstupních zjištění pro analýzu.

3.1 Báze strojového učení

Problematika strojového učení vychází z rozvoje umělé inteligence na počátku šedesátých let, kdy obor získával pevnější rámec. Historická forma byla oproti současné digitální založena především na rozboru statistických fenoménů, kdy její zaměření vedlo k následnému poklesu zájmu o problematiku. Následné období takzvané „zimy umělé inteligence“ značně zpomalilo postup vývoje prakticky okamžitě po implementaci Baysiánských metod. Za současnou formu strojového učení vdčíme období devadesátých let dvacátého století, kdy došlo k opětovnému oživení oboru. S příchodem objemných datasetů, relativně dostupné a výkonné výpočetní techniky se vlastní tvorba neuronových sítí posunula mimo rovinu koncipování potenciálních přístupů (Miikkulainen et al. 2019). Následný nárůst výkonnosti výpočetní techniky umožnil další škálování algoritmů a jejich implementaci do oborů strojového vidění či rozpoznání řeči.

Rámcově se jedná o počítačový program, který učení získává zkušenosti zařazené do předurčené třídy, které následně lze měřit na základě vlastní výkonnosti v jejich reprodukci (Goodfellow et al. 2016). Bází samotného nabírání zkušeností je praktická prezentace vybraných dat, kdy v případě strojového vidění dochází k využití předem determinovaného počtu filtrů. Jednotlivé filtrové vrstvy následně odesílají vlastní výstupy následným vrstvám neuronové sítě, která v průběhu času a opakovaných analýz postupně získává generalizační zkušenost s jednotlivými objekty. Vzhledem k růstu rozsahu sítí lze procesy provádět primárně v rámci výkonné výpočetní techniky, jejíž simulační schopnosti urychleně procesují jednotlivé elementy dat neuronovou sítí v determinovaném čase takzvaných epoch (LeCun et al. 2015).

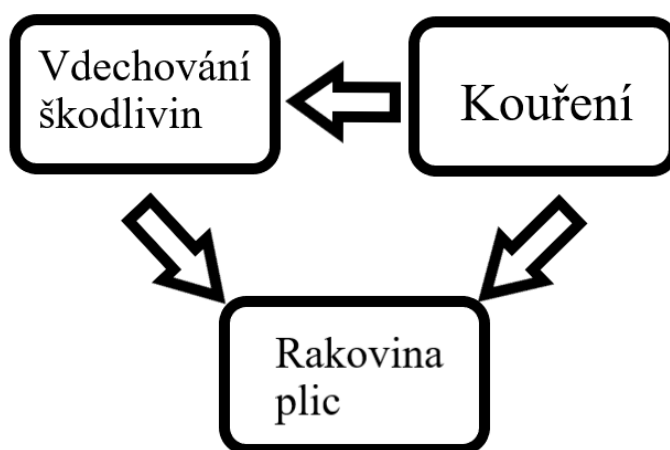
Publikace tvořící jádro oboru pečlivě popisují potřebné prvky pro formulaci výzkumu užívaného strojového učení, tedy: dataset, hodnotovou funkci, výpočetní model a optimalizační proceduru (Goodfellow 2016). Jednotlivé aspekty lze z matematického hlediska popsat poměrně jednoduchými funkcemi, například výpočetní model může být lineární či nelineární, ovšem jejich společné fungování zdatelně přesahuje pravděpodobnostní statistické přístupy starších modelů. Celkový posun oboru od probabilistického výpočtu jasně determinuje potřeba přesných výstupů, kdy soudobá implementace algoritmů do báze sociálních sítí vyžaduje přesnost zodpovědnou za ekonomické výsledky společností. Přesto, že je strojové učení považováno za revoluční přístup k problematice, jeho základ pochází z praktické inspirace neurologií a fungováním vnitřních struktur lidského mozku (Goodfellow 2016).

Trénink neuronových sítí lze dle LeCun et al. (2015) realizovat dvěma fundamentálními postupy, trénováním s dohledem či bez něj. Za trénink s dohledem považujeme implementaci algoritmu na dataset disponující zcela jasnou informační rovinou ve formě anotací, kdy jsou jednotlivé objekty prezentovány, jakožto prvky pevně daných kategorií. Ač článek považuje postup s dohledem do budoucna za méně efektivní, tak tvorba modelů bez dohledu při tréninku v současnosti značně zaostává za modely s dohledem vzhledem k vyšší výpočetní náročnosti a nižší přesnosti detekce. Pro budoucí vývoj je možné spekulovat o dalším zlepšení výstupů tréninku bez lidského dohledu, vzhledem k jeho potenciálnímu šetření práce nutné pro formulování úspěšného modelu. Z obecného hlediska lze přípravě dat přiřadit až 80 % času nutného pro tvorbu modelu, vzhledem k realitě náročnosti anotování dat.

Tvorba anotací, ačkoliv zcela zásadní pro fungování strojového vidění a širšího kontextu strojového učení, zůstává primárně manuální činností pro lidské operátory. S postupem technologického pokroku vznikla řada nástrojů zjednodušujících tento proces (Hansen et al. 2021), ovšem jeho úplné vyřazení z práce s modely se v současnosti zdá být nepravděpodobné. Zásadním problémem tvorby anotací zůstává i jejich potenciální chybovost, kdy možnosti kontroly výstupů u objemnějších datasetu nenaplňují prvky reálnosti. Samotná chybovost popisu dat může výsledně kompromitovat neuronovou síť a v období adversních dat zásadně zredukovat její spolehlivost (Miceli et al. 2020). V rámci spolehlivosti je za určitou slabinu obou modelů možné považovat jejich značnou absenci generalizace, která v nekontrolovaném prostředí může vyústit až v poměrně razantní nespolehlivost modelu (Goodfellow et al. 2016).

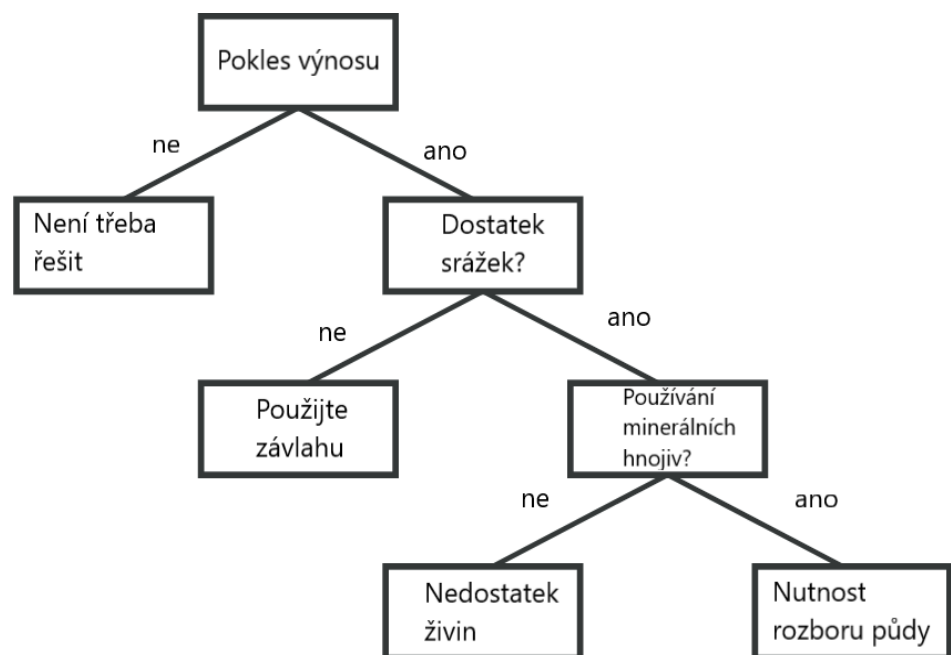
3.1.1 Metody v rámci strojového učení

Z hlediska jednotlivých metod strojové učení je disciplína rozdělena dle komplexnosti a pokročilosti na pět základních přístupů k problematice zpracování trénovaného algoritmu. Jako první z moderních přístupů k strojovému učení je použití Bayesiánských sítí, které jsou v podstatě necyklickými sadami nezávislých proměnných a jejich podmíněnými vztahy (Heckerman & Wellman 1995). Principem modelů je pravděpodobnostní výstup založený na vztahu vstupních sledovaných dat s jejich potenciálními zdroji (viz Obrázek 1). Jedním z nejčastějších příkladů Bayesiánských sítí jsou pravděpodobnostní vztahy mezi nemocemi a jejich příznaky, kdy výstupem modelu je pravděpodobnost diagnóz na základě vstupních dat o příznacích. Dalším metodickým přístupem k strojovému učení je využívání regresní analýzy, která je v jádru primárně statistickou metodou. Nejčastějším nástrojem analýzy je lineární regrese, která prakticky kopíruje nejpřesnější možný průnik datových vstupů, na jehož základě produkuje pravděpodobnostní výstupy (Recenti et al. 2019). Podobně jako Bayesiánské sítě se metody používá často v medicínském prostředí, kde procesy často podléhají lineární logice snadno zpracovatelné regresí.



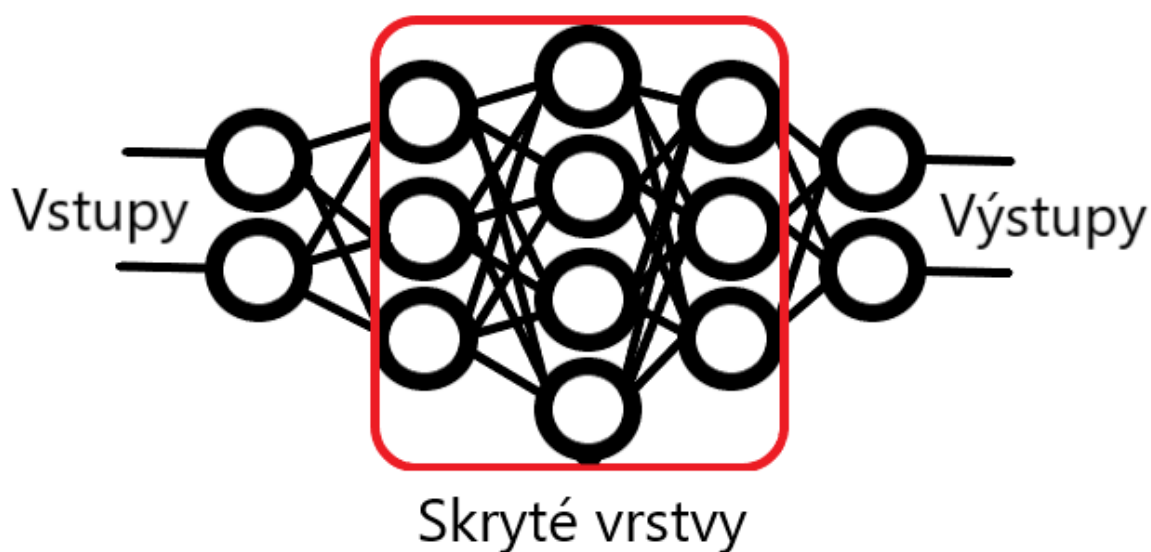
Obrázek 1 Jednoduchá Bayesiánská síť

Alternativou k primárně pravděpodobnostním metodám je použití metody podpůrných vektorů (Support Vector Machines). Základem přístupu je v podstatě binární rozdělení prezentovaných dat, které jsou posléze modelem evaluovány v rámci tréninku. Výstupem algoritmu je predikce příslušnosti vstupních dat k jedné z kategorií, kdy je model prakticky rezistentní proti výskytu chyby I. typu (Ahmed et al. 2012). Druhým z nepravděpodobnostních přístupů k strojovému učení je systém stromových struktur (decision trees). Model funguje skrze zhodnocení vstupních dat, která jsou reprezentována kmenem a větvemi a jejich přiřazení k cílové kategorii reprezentované listy. Rámcově stromová struktura posuzuje vstupní data na bázi predeterminovaných kritérií, která postupně směřují k výstupním kategoriím na konci sítě, přičemž u jednoduchých forem struktury disponují výstupními kategoriemi i úvodní rozvětvení (viz Obrázek 2).



Obrázek 2 Příklad jednoduché stromové struktury

Pokročilým přístupem k strojovému učení je používání umělých neuronových sítí, jejichž základní logika pramení ze snahy o imitaci fungování lidského mozku. Ačkoliv současné porozumění rozsahu fungování nervové soustavy lidského těla zůstává neúplné, inspirace fungováním vzájemné interakce nervových buněk se projevila jako efektivní (LeCun et al. 2015). Algoritmus neuronové sítě vysílá vstupní informace do tzv. „skrytých vrstev“, kde dochází k interakci dat s jednotlivými neurony, které zpracovávají přijaté informace a pošlou vlastní výstup do napojených neuronů další vrstvy (viz Obrázek 3). Výhodou přístupu je jeho potenciální škálovatelnost, kdy s narůstajícím počtem vrstev a jednotlivých neuronů může algoritmus procesovat závratné množství dat. Přesto, že se v současnosti technologicky jedná o přístup s nejšířší schopností zobecnění dat, jsou schopnosti algoritmů mimo kontrolované prostředí do značné míry náchylné k nepředpokládaným vstupům (Goodfellow et al. 2016).



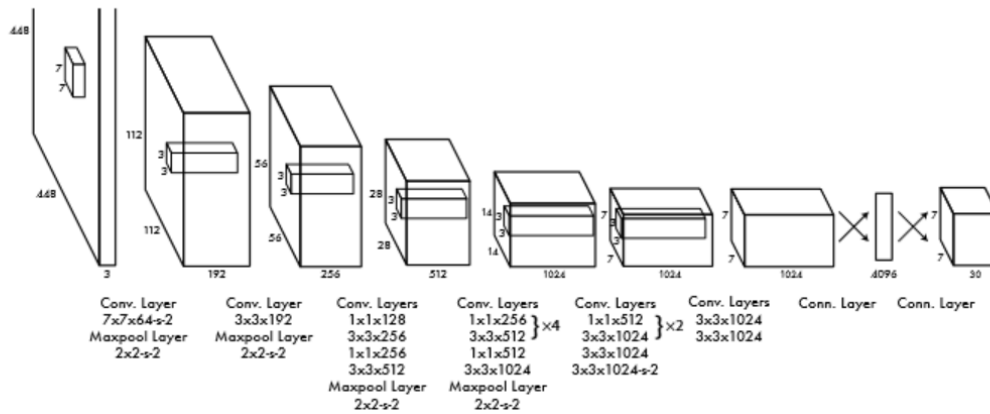
Obrázek 3 Umělá neuronová síť

Pro úspěšné zpracování vizuálních dat se v rámci umělých neuronových sítí vytvořila specifická subdisciplína používající konvoluční neuronové sítě. Odlišnost konvolučních sítí od běžných umělých neuronových sítí spočívá v jejich architektuře sdílených vah v konvolučních filtrech, které převádí jednotlivé součásti vstupních dat na prvkové mapy (Farooq et al. 2019). Základem fungování konvoluční architektury je plné propojení neuronů se všemi neurony následující vrstvy, které při správném nastavení parametrů tvoří filtr zaznamenávající vzory ve vstupních datech. Kombinace rekurentních vzorů je jádrem detekce objektů na vizuálních datech, přičemž skrze jednotlivé parametry konvolučních filtrů funguje rozvoj zobecnění elementů detekce. Celkově lze konvoluční síť chápat jako sadu na sebe navazujících filtrů rozdělujících vstupní obraz na jednotlivé prvky, z nichž se následně skládají objekty a eliminují necílové segmenty dat.

3.1.2 Neuronové sítě YOLO

Jako srovnávací prvek pro vlastní model byl v rámci tvorby analýzy vybrán model YOLOv3. Z obecného hlediska se model prosadil, jakožto jeden z nejrychlejších v rámci zpracování vstupních dat, kdy již první verze z roku 2016 překonávala většinu svých konkurentů. Verze 1 byla schopna zpracovávat v odlehčené formě data rychlostí 155 snímků za sekundu a v plné procesní formě přibližně 45 snímků za sekundu (Redmon et al. 2016). Současná verze 3 je považována za lépe vybalancovanou, přičemž ve srovnání dosahuje až 3,8 krát vyšší výkonnosti než RetinaNet při zachování obdobné přesnosti detekce (Redmon & Farhadi 2018). Vzhledem k jeho rychlosti bylo možné považovat model za tzv. „real time“, tedy schopný detekce v dostatečné rychlosti, aby pro lidského pozorovatele působila okamžitě. Celkově se jedná o v současnosti jeden z nejpoužívanějších modelů pro detekci objektů ze snímků či záznamů v jak výzkumném, tak praktickém prostředí.

Model jako takový byl vybudován na logice regresního přístupu k prostorově odděleným boxům nesoucí jejich asociovanou pravděpodobnost k jedné z tříd. Došlo tedy k odchylce od konvenčního pohledu na individuální trénink jednotlivých komponentů, které byly nahrazeny jednou primární regresní problematikou (Redmon et al. 2016). Celkově model benefitoval z vlastní konstrukce zaměřené na eliminaci přebytečných kroků, kdy v průběhu procesování analyzuje vstupní snímky jako celek, nikoliv jako mozaiku individuálních segmentů. V průběhu procesování rozdělil systém obraz na mřížku $S \times S$, která slouží jako báze pro alokování predikcí. Jednotlivé buňky v mřížce obsahují B vazných boxů, skrze které popisuje jistotu vlastní predikce. Systém následně přidělil v rámci jednotlivých boxů poziční hodnoty ve vztahu k mřížce a celku snímku, doplněné o hodnotu pravděpodobnosti vlastní predikce. Finálním krokem analýzy bylo přidělení hodnoty C reflektující kondiční pravděpodobnost jednotlivých tříd. Dohromady systém YOLO vyprodukoval tenzor $S \times S \times (B * 5 + C)$ zodpovědný za predikční schopnosti algoritmu (Redmon et al. 2016).



Obrázek 4 Architektura modelu (Redmon et al. 2016)

V rámci poslední verze modelu byla využita architektura založená na původní verzi systému (viz Obrázek 4), díky níž je algoritmus schopen dosahovat výše zmíněných výkonů. Při srovnání s konvenčními modely byl systém YOLOv3 svou architekturou poměrně jedinečný díky eliminaci veškeré normalizace a prioritizace přímého průchodu dat (Redmon & Farhadi 2018). Z hlediska modernizace třetí verze došlo oproti verzi první k nahrazení dvou plně propojených vrstev (viz Obrázek 4) ve výstupových vrstvách neuronové sítě modernější kombinací vrstev (viz Obrázek 5), která dále zvýšila výkonnost modelu jako celku. Hlavním posunem nejmodernější verze pojmenované Darknet-53 bylo oproti verzi základní zvýšení počtu konvolučních vrstev z 24 na 53 (Redmon & Farhadi 2018).

| | Type | Filters | Size | Output |
|----|---------------|---------|-----------|-----------|
| | Convolutional | 32 | 3 × 3 | 256 × 256 |
| | Convolutional | 64 | 3 × 3 / 2 | 128 × 128 |
| 1x | Convolutional | 32 | 1 × 1 | |
| | Convolutional | 64 | 3 × 3 | |
| | Residual | | | 128 × 128 |
| | Convolutional | 128 | 3 × 3 / 2 | 64 × 64 |
| 2x | Convolutional | 64 | 1 × 1 | |
| | Convolutional | 128 | 3 × 3 | |
| | Residual | | | 64 × 64 |
| | Convolutional | 256 | 3 × 3 / 2 | 32 × 32 |
| 8x | Convolutional | 128 | 1 × 1 | |
| | Convolutional | 256 | 3 × 3 | |
| | Residual | | | 32 × 32 |
| | Convolutional | 512 | 3 × 3 / 2 | 16 × 16 |
| 8x | Convolutional | 256 | 1 × 1 | |
| | Convolutional | 512 | 3 × 3 | |
| | Residual | | | 16 × 16 |
| | Convolutional | 1024 | 3 × 3 / 2 | 8 × 8 |
| 4x | Convolutional | 512 | 1 × 1 | |
| | Convolutional | 1024 | 3 × 3 | |
| | Residual | | | 8 × 8 |
| | Avgpool | | Global | |
| | Connected | | 1000 | |
| | Softmax | | | |

Obrázek 5 Modernizovaná architektura modelu (Redmon & Farhadi 2018)

3.2 Koncepční prvky

3.2.1 Obecně koncepční články

Za úplný počátek koncepce vizuálně řízené aplikace chemických či mechanických zásahů proti plevelným rostlinám můžeme historicky považovat prvotní rozvoj precizního zemědělství. Původní jádro oboru se soustředilo na zlepšení přesnosti zásahů skrze využití geolokačních dat a jejich zapojení do zpřesnění pohybu pracovních prostředků po pozemku v rámci daných operací (Davis et al. 1998). Rámcově koncepce fungovala s rozvojem GPS, které pro civilní užití umožňovalo relativně znatelné zpřesnění řady procesů. Ačkoliv další rozvoj kombinování družic a pozemních geolokačních bodů umožnil postupné zpřesnění polohy až k hranici jednotek centimetrů, chemické aplikace byly stále vedeny plošně. Dalším logickým krokem v rozvoji oboru tedy byla kontinuální precizace jednotlivých kroků.

Potenciálním kandidátem pro řešení dalšího zpřesnění zásahů byl obor umělé inteligence, který podporoval precizaci a autonomizaci oboru (Davis et al. 1998). Reálných výsledků ovšem bylo technicky možné na rovině jednotlivých rostlin dosáhnout až s nástupem technologie strojového učení. Rozvoj disciplíny s postupnou adaptací technologie na jednotlivé podobory zahrnující strojové čtení (později strojové vidění) vedl postupem času k implementaci technologie na širokou škálu specializací. Pro detekci plevelů se stala zásadním právě rovina strojového vidění, která umožňovala potenciální odklon od plošných zásahů (Gebbers & Adamchuk 2010). Navzdory existujícímu potenciálu ovšem technologie stála na počátku náročné cesty s řadou teoreticko-systematických i praktických překážek.

Za jedny z prvních problémů k řešení v rámci disciplíny patřily dle McCarthy et al. (2010) otázky sběru a segmentace dat. Článek vedle identifikace potenciálních datových platforem vyzdvihl i fundamentální problematiku dělení implementace na vnitřní (kontrolované) a vnější (nekontrolované) prostředí aplikace. Text systematicky analyzuje potenciál a nedostatky tří hlavních nástrojů pro sběr dat, tedy RGB snímků, stereografického 3D zobrazení, multispektrálních obrazů a pasivního hloubkového měření (McCarthy et al. 2010). Jednotlivé platformy jsou posuzovány nejen na základě vlastností získaných dat, ovšem i s ohledem na typ prostředí, v které jsou nasazovány, tedy výše zmíněné dělení na vnitřní a vnější prostor. Celkově se metodicky jedná o kvalitní bázi pro další rozvoj výzkumu v oboru.

V obdobném smyslu se s desetiletým odstupem k problematice postavil kolektiv autorů, jehož primárním cílem byla identifikace potenciálního užití algoritmů strojového učení v agrární sféře a jeho operacionalizace (Sharma et al. 2020). Článek systematicky analyzuje jednotlivé prvky potravinové produkce a přiřazuje k nim relevantní publikace. Cílem systematického přehledu je vedle představení potenciální řešení v rámci jednotlivých oblastí i jejich provázání napříč oborem potravinářství. Koncepce strojového učení tvoří maximalistický systém, kdy využívání datových platforem může při koordinaci více prvků oboru značně omezit ztráty a zvýšit výkonnost. Srovnání jednotlivých výzkumů v oboru dále umožňuje článku tvorbu předpokladů pro další vývoj ve výsledné přesnosti detekce zkoumaných dat.

Vzhledem k posunu výzkumného paradigma je článek veskrze směřován k primárně udržitelné formě potravinové produkce, která podobně jako obor strojového učení patří mezi nejzásadněji rostoucí segmenty. Význam detekce plevelů pro udržitelné směřování agrární sféry tkví v kombinaci preciznějších chemických zásahů s limitovaným přesahem na prostředí a zvýšení přesnosti a účinnosti mechanických zásahů. Oba koncipované přístupy mohou znatelně zvýšit ekonomickou produktivitu agrární sféry a dále zlepšit její úspěšnost při správě a udržování antropocenóz. Článek společně s představením situace v oboru poukazuje i na technologické posuny v rámci analytických operací, kdy přehled dostupného výzkumu jasně demonstrovuje prioritizaci hlubokých neuronových sítí oproti dříve formulovaným metodám (Sharma et al. 2020).

3.2.2 Publikace koncipující vlastní projektové řešení

Ačkoliv se disciplína prakticky vyvíjela již od počátku nového tisíciletí v krátké návaznosti na prvotní pokroky samotného strojového učení, úvodní dekáda nabídla prvních komerčně aplikovaných projektů (Gerhards et al, 2002). O zásadní pokrok v oblasti detekce plevelných rostlin na bázi vizuálních dat se na závěru období dále podílel tým zodpovědný za konferenční příspěvek z roku 2010, který se zabýval detekcí plevelných druhů ve vzcházejících porostech kukuřice (Chen et al. 2010). Přestože je v dnešní době tehdejší primárně probablistický model relativně nedostačující, tak metodické přístupy k problematice lze implementovat i na hlubokých neuronových sítích. Publikace řeší problematiku oddělení rostlinných částí od pozadí ve formě půdy či posklizňových zbytků s úspěšností v okolí 95 procent, kdy přesnost detekce kukuřice dosahuje přibližně 92,5 % (Chen et al. 2010). Implementace limitovaného datasetu, jakožto jádro modelu, umožňuje autorům poměrně vysokou úspěšnost v rámci vybraných dat. Celkově ovšem vzhledem k vlivům prostředí je aplikovatelnost modelu do praxe poměrně limitovaná, především díky relativně malé robustnosti a špatné škálovatelnosti probablistických metod.

Pětileté období navazující na úvodní periodu implementací technologie do agrární sféry se vyznačovalo rozvojem četných přístupů k strojovému učení. Mezi nejúspěšnější patřily LDA (lineární diskriminační analýza), DT (decision trees) a SVM (support vector machines), které se prosadily například v předpovědi výnosů či detekci symptomů chorob plodin (Liakos et al. 2018). Příklady testování SVM na vybraných digitálních datech dokázaly v rámci specifických podmínek zajistit i relativně vysokou spolehlivost detekcí s přesahem 97 % (Ahmed et al. 2012). Za poměrně silnou stránku můžeme retrospektivně považovat i poměrně stabilní absenci chyb I. typu, které jsou v současnosti reálným aspektem většiny používaných přístupů. Obecně ovšem technologie narážela na problematickou škálovatelnost jak v otázce velikosti datasetu, tak v architektuře použitých modelů.

Primárním elementem rozvoje výše zmíněných přístupů k implementaci strojového učení do oboru byl postupný nárůst významu udržitelnosti produkce spojené s neochotou společnosti konzumovat potraviny obsahující residua pesticidů. Dalším elementem v kontextu environmentálních snah společnosti byla hlediska ochrany vodních zdrojů či biodiverzity. Jednotlivé úspěšné implementace ve formě předpovědi výnosů či detekci chorob společně s postupem času našly praktické uplatnění v rámci konceptu precizního zemědělství. Jako předpokládaný důvod pokračující snahy o další rozvoj vizuální detekce se projevil zásadní dlouhodobý význam plevelných rostlin na výstupní výnosy produkce (Gharde et al. 2018), jehož řešení převzala řada výzkumníků jako svou prioritu ve vztahu ke společenským potřebám. Celkově jednotlivé prvky tvoří významné elementy rámce precizního zemědělství, kde dopomáhají v dosahování maximalistických přístupů k zpracování výstupních dat (Gebbers & Adamchuk 2010).

Navzdory postupnému vývoji a často i úspěšné implementaci napříč širším oborem se v problematice detekce plevelů zásadně prosadil až nástup hlubokých neuronových sítí (Sharma et al. 2020). Důvody technologického vývoje lze hledat v řadě faktorů, kdy v první řadě hluboké neuronové sítě dlouhodobě přinášely komparativně nejpřesnější výsledky detekce. Alternativním přístupem můžeme na problematiku nahlížet ve světle technických prvků, kdy pro detekci plevelných rostlin zůstává zásadní rozpoznání tvarů naproti barevnému spektru nutnému pro správné zhodnocení výskytu chorob (Liakos et al. 2018). S dalším postupem výkonnosti architektur hlubokých neuronových sítí většina ostatních zmíněných přístupů postupně poklesla na významu.

Efektivní implementace konvolučních neuronových sítí v širším portfoliu sledovaných tříd dle Dyrmann et al. (2016) umožnila dosažení úspěšnosti přesahující 86 %. Ve srovnání s vlastními předchůdci byl algoritmus užitý pro detekci 22 druhů rostlin. Většina detekovaných objektů byla tvořena plevelnými rostlinami, kdy se jednalo o nejvýznamnější dvouděložné plevele v Dánsku a uniformní identifikátor jednoděložných plevelných rostlin (Dyrmann et al. 2016). Samotný dataset výzkumu přesahoval svým finálním rozsahem úroveň 10000 objektů, které v rámci 22 tříd disponovaly relativně proporcionálním zastoupením. Článek rámcově vyzdvihuje význam znásobení úvodního objemu dat skrze jejich rotační úpravy, přičemž výsledný dataset zlepšuje finální robustnost modelu k variabilním vstupům.

V rámci metodiky článek definuje rovnoměrnost tréninkových dat, kdy nutnost obdobné pixelové velikosti objektů eliminuje potenciální falešnou korelaci tříd objektů na základě jejich rozměrů (Dyrmann et al. 2016). Ačkoliv současné formy strojového učení dokáží vliv velikostní variability značně omezit, problematika kompatibilní datové variability zůstává pro obor zásadní. Závěrečným výstupem je vize mobilního nástroje pro mapování rozsahu zaplevelení, který umožňuje zemědělcům efektivní rozhodování v rámci aplikovatelných látek. Variabilní aplikace herbicidů přizpůsobená aktuálnímu výskytu plevelných druhů podporuje jejich cílenou eliminaci a snížení vstupů do prostředí ve vztahu k necílovým organismům, kdy článek dobou svého vydání reflektuje postupný posun výzkumného paradigma z ekonomické efektivity na environmentální udržitelnost.

Dalším krokem v užívání konvolučních neuronových sítí bylo rozšíření spektrálních vlastností snímků. Ač samostatné používání RGB snímků nabízí poměrně široké možnosti, jeho doplnění o NIR kanál umožnilo zásadní zvýšení úspěšnosti detekce rostlin vůči pozadí (Potena et al. 2017). Společně s narůstající celkovou přesností klasifikace stoupala i úspěšnost v rámci rozpoznání jednotlivých druhů, způsobená především rostoucími rozměry tréninkových dat, ke kterému vedl i nárůst užívání komerčních dronů pro sběr dat (Dos Santos Ferreira et al. 2017; Kerkech et al. 2018). Ačkoliv tréninková data většiny výzkumů užívajících drony obsahovala řádově stovky snímků, rozsah jednotlivých objektů dodávaných algoritmu přesahoval jednotky tisíců. Výsledný nárůst dat v zásadě napomohl k robustnějším výsledkům a souběžnému zpřesnění detekcí.

S rozvojem dálkového průzkumu v podobě užívání bezpilotních snímkových platform došlo k fundamentálnímu rozdělení na dvě subdisciplíny. Na jedné straně plošné záběry ploch z leteckých snímků zásadně zvyšovaly výkonnost platform a objem tréninkových dat (Liakos et al. 2018), na straně druhé ovšem za zvýšenou výkonnost platily snížením prostorového rozlišení snímků ve vazbě na jednotlivé objekty. Opačným směrem se vydala subdisciplína zaměřená na primárně pozemní snímkovací platformy, jejichž primárním cílem bylo zaměřením na detailní snímky objektů umožňující precizní detekci jednotlivých druhů (Wang et al. 2019). Zásadním rozdílem obou subdisciplín se stalo i spektrální rozlišení snímků, kdy u platform dálkového průzkumu převládlo užití hyperspektrálních kamer v kontrastu s rozmachem RGB a multispektrálních kamer u pozemních platform (Liakos et al. 2018).

V současnosti je již u výzkumů v obou subdisciplínách možné sledovat poměrně vysokou úspěšnost detekce jednotlivých objektů. Rámcově se díky rozvoji hlubokých neuronových sítí daří radě týmů dosahovat hodnot přesahujících 90 % úspěšnosti klasifikace bez ohledu na formu dat, tedy RGB, multispektrální, pozemní či dálkově snímaná. Dle Farooq et al. (2019) lze úpravou datasetu hyperspektrálních snímků dosáhnout až 96 % úspěšnosti v rámci zkoumané selekce druhů, která ovšem zůstává relativně limitovaná. Příčinou limitů pro datový rozsah analýz je na jedné straně náročnost přípravy snímků pro trénink modelů a na straně druhé i stoupající výpočetní zátěž při použití komplexnějších podmínek.

Finálním problémem úspěšné detekce je i čas, který je nutný pro úspěšné provedení výpočetní operace, kdy současné studie dosahují hodnot v okolí 57 milisekund (Alam et al. 2020). Ačkoliv se časový rámec prezentovaný autory může zdát dostatečným, praktické využití technologie vyžaduje zásadní zrychlení. Rámcově se tedy objevují projekty zaměřené na odlehčení modelů, skrze které se časová náročnost výpočetních operací v praxi znatelně snižuje (Abouzahir et al. 2022). Trvalým problémem příklonu k rychlosti však vzhledem k odlehčení architektury algoritmu zůstává realita poklesu přesnosti, jejíž navýšení zůstává otázkou dalšího rozvoje disciplíny a vylepšení jednotlivých nástrojů.

Konkluzivně je možné pohlížet na celkový posun v oboru dvojicí perspektiv. V první řadě je zde dlouhodobý nárůst úspěšnosti detekce (Sharma et al. 2020), kterým ovšem kompletně nekoresponduje s nárůstem objemu dat a množstvím sledovaných tříd. V logice škálování technologie je zastoupení co nejvyššího počtu objektů a tříd zásadní pro spolehlivost a robustnost modelů. Výsledné škálování tedy ve světle většiny výzkumů zaostává za hranicí nutnou pro úspěšné zařazení technologie do praxe. Druhou rovinnou perspektivy je zařazování soudobých inovací do výzkumu, kdy rozličné datové platformy doplněné o lehčí a rychlejší modely umožňují potenciální rozvoj výzkumu. Přesto, že druhá perspektiva dává značnou naději pro rozvoj oboru, zůstává potřeba ruční přípravy dat nejzásadnějším omezením jeho zásadnějšího rozvoje, kde ovšem přibývá snah o tvorbu nástrojů k zjednodušení procesu (Hansen et al. 2021).

3.3 Empirický potenciál pro praktickou implementaci

Jádrem výzkumné tvorby modelů schopných detekce plevelných rostlin zůstává problematika jejich implementace do praktických polních podmínek. Rámcově se v koncepci zcela jednoznačně projevuje otázka ekonomického dopadu zařazení technologie do praxe, která v současnosti významně doplňuje základní environmentální motivaci (Pardey et al. 2010). Vedle zásadního významu ekonomiky na rozvoj technologií zůstává v případě strojového učení i neopomenutelná rovina robustnosti modelu a jeho schopnost odolávat adversním podmínkám nekontrolovaného prostředí (Madry et al. 2017). Reálné zařazení technologie do oborové praxe je možné pouze skrze úspěch v obou fundamentálních otázkách, neboť bez poptávky a fungujícího nástroje technologie zůstává pouze prvkem akademického prostředí.

3.3.1 Ekonomický efekt

Navzdory nesčetným deklaracím a úmluvám zaměřeným na inovace zůstává pro zásadní majoritu technologických změn hlavním faktorem jejich ekonomický přínos. V čistém kontextu šetření chemických prostředků může použití strojového vidění působit, jako ekonomicky jednoznačně výhodná technologie. Tento argument stvrzují i výstupní hodnoty z výzkumů zaměřených na rozsah úspory při použití aplikačních map při potlačování výskytu plevelných rostlin pod stanovený práh škodlivosti (Timmermann et al. 2003). Realitu ekonomické výhodnosti ovšem determinuje kombinace řady faktorů, které přímo či nepřímo vyplývají z vnitřní dynamiky agrární sféry spojené s jejím dlouhodobě komplikovaným vztahem k inovacím (Guerin & Guerin 1994). Celkově lze ekonomický význam pro praxi rozdělit do třech základních rovin, tedy tržní, technologické a podnikové.

Tržní prostředí má vzhledem k jeho roli zásadní vliv na technologický rozvoj. Problémem jeho vlivu ovšem zůstává čirá ekonomická motivace, kterou lze v případě strojového vidění a detekce plevelných rostlin uplatnit na otázku globální potřeby technologie. Ačkoliv je využívání chemické ochrany plodin globálním fenoménem, v rámci jednotlivých regionů existují poměrně zásadní rozdíly v rozsahu a formě. Ve světle zásadních rozdílností je potřeba redukce aplikovaných látek zcela jednoznačně odlišná, přičemž s výjimkou Evropské unie dochází ke stálému globálnímu nárůstu významu herbicidů (Gianessi 2013). Logikou rozvoje herbicidních zásahů je zcela jednoznačně jejich cena, kdy díky rozsahu chemického průmyslu dosahují na nabídku chemických látek i rozvojové země.

Ačkoliv jsou legislativně motivováni k redukci herbicidních zásahů prakticky pouze evropští zemědělci, jejich tržní význam není jediným důvodem pro nadnárodní korporace k investici do technologie. Ekonomický význam Evropského agrárního trhu zcela jednoznačně reflektuje jeho cílený i pasivní vliv na světové zemědělství (Prins et al. 2011). Vedle významu evropského zemědělství se ovšem v dnešní době projevuje kombinace dvou faktorů, které by pro globální trh jednoznačně navýšily význam strojového vidění. Na jedné straně se vlivem rostoucích cen energií do určité míry omezuje ekonomická efektivita chemických zásahů, což se v případě rozvojových zemí projevuje užitím levnějších a často rizikovějších přípravků. Na straně druhé je v současné době mezi nejsilnějšími trendy odklon od neselektivních chemických zásahů, jejichž negativní důsledky pro životní prostředí slouží jako kritický bod společenské perspektivy (Camargo et al. 2020). Oba tyto trendy lze do určité míry zmírnit pomocí širší implementace strojového vidění, skrze něž by bylo možné zvýšit cílenost herbicidních zásahů a současně ušetřit a zbezpečnit chemickou rovinu problematiky.

Z ekonomického hlediska je technologická perspektiva poměrně komplikovaným faktorem, který dokáže fundamentálně ovlivnit výsledný formát implementovaného přístupu. Základem technologické roviny je v zásadě cenový poměr nutné vstupní investice do vývoje s pravděpodobností a předběžnou dobou její návratnosti, celkově tedy inovace představují do určité míry rizikovou investici (Lazonick & Mazzucato 2013). Implementace strojového vidění do praxe by v rovině inovační investic vyžadovala kombinaci kamer disponujících relativně vysokým rozlišením, doplněných o přenosovou soustavu napojenou na palubní počítač schopný zpracování neuronové sítě (Feng et al. 2019). Ačkoliv se z historického hlediska cena kamer dlouhodobě snižuje, kamery dostatečného rozlišení v RGB formátu zůstávají pro širší rozlišení stále poměrně drahé. Jako potenciální alternativa by se díky své nízké ceně nabízel formát použitý v rámci práce, tedy infračervené kamery (Luo et al. 2010). Problém jejich samostatného užití ovšem zůstává limitace přesnosti snímků vlivem vnějšího prostředí.

Vedle ceny kamer zůstává poměrně problematickým aspektem i výpočetní technika nutná pro provoz neuronové sítě zodpovědné za detekci objektů zachycených kamerou. Navzdory potenciálnímu odlehčení prakticky používaných modelů a přizpůsobení potřebných komponentů (Feng et al. 2019) je potřebný hardware v současnosti poměrně složitě dostupný díky potížím v dodavatelském řetězci a produkci mikročipů (Kamasa 2021). Případná saturace nově vzniklého trhu by v současné situaci vytvořila další růst cen doprovázený chronickými nedostatky náhradních dílů, čímž by širší implementace technologie narážela na časté periody obecné stagnace. Konkluzivně lze obtíže v rámci technologické roviny překonat primárně na evropském trhu a americkém trhu, které ve srovnání se zbytkem světa disponují potřebným kapitálem, motivací a významem (Pardey et al. 2010). V kontextu širšího světa lze ve scénáři implementace technologie počítat s dlouhodobějším časovým horizontem, kde cena jednotlivých elementů vzhledem k rozvoji dodávek do prvního světa postupně klesá.

Finálním ekonomickým aspektem zůstává otázka podnikové úrovně, která se v práci zaměřuje na evropské prostředí vzhledem k výše zmíněné problematice globálního trhu. Kontext společné zemědělské politiky Evropské unie zcela jednoznačně vyzdvihuje význam snížení potřeby chemických zásahů, problém ovšem nastává u dimenzí podniků (Antoniuk et al. 2017) a reality směřování podpory přímo k inovacím. Navzdory dlouhodobé snaze společné zemědělské politiky jsou ve vztahu k rozvoji stále nejvíce náchylné malé a střední podniky, které by s financováním technologie měly značné obtíže. Jako potenciální řešení situace se nabízí další rozšíření podpory kýmých podniků, ovšem úspěšnost politicky podporované inovace zůstává poměrně problematická (Lazonick & Mazzucato 2013). Jeho slabinou je absence zásadních mechanismů determinujících, kterou technologií podpořit, doplněná o relativně nízkou motivaci k jejich tvorbě a financování. Celkově je tedy v rámci podpory ze strany Evropské unie inovační přístup přehlížen a jeho primárními tvůrci zůstávají korporace a univerzity (Gerhards et al. 2022).

Pro jednotlivé podniky je vedle snahy o následování pravidel společné zemědělské politiky významnou otázkou i případný poměr ušetřených prostředků vzhledem k investici. Při předpokladu realistické hranice 50 % ušetřených prostředků (Timmermann et al. 2003) zůstává jednorázová investice vzhledem k vývoji cen inovací ekonomicky výhodnou ve vztahu k rostoucímu objemu produkce (Pardey et al. 2010). Teoreticky je možné předpokládat snahu středních a velkých podniků o využití jejich rozsahu pro maximalizaci zisku z inovace. Celkově lze dojít k závěru, že praktická implementace technologie je z ekonomického hlediska možná, ovšem podléhá řadě problematických faktorů, které mohou proces značně zkomplikovat.

3.3.2 Odolnost k nekontrolovanému prostředí

Ekonomický element předpokládaného výkonu pracuje primárně s výstupními daty současného výzkumu, který ovšem do značné míry neobsahuje problematiku nekontrolovaného prostředí implementace. V současnosti za jednu ze zásadních slabin technologie strojového učení považujeme jeho náchylnost k vlivům nekontrolovaného prostředí. Přesto, že modely často dosahují poměrně vysokých přesností v identifikaci dat, výskyt adversních podmínek může poměrně snadno výrazně narušit finální výstup (Madry et al. 2017). Největším rizikem adversních dat je jejich snadná přehlédnutelnost, kdy pro člověka nerozpoznatelná odlišnost dokáže radikálně narušit spolehlivost neuronové sítě. Problematika adversních dat je navíc pro neuronové sítě kumulativní s časem, kdy data mohou fungování algoritmu narušit jak při tréninku, tak během implementace (Huang et al. 2011). Jádrem zranitelnosti zůstává u strojového učení jeho absence vnímání obecnějších konceptů, které v lidské perspektivě umožňují komplexnější asociace (Madry et al. 2017).

Ačkoliv je obor adversních dat primárně zaměřen na lidské oponenty, které lze jen velmi těžko očekávat v agrární sféře, příroda zcela bez pochyb disponuje nástroji schopnými obdobných výsledků. Současné modely dosahují poměrně stabilně zlepšujících se výsledků ve vztahu k detekci plevelných rostlin, ovšem značný počet velmi přesných modelů dosahuje nižších výsledků při inkorporaci tříd plodin (Liakos et al. 2018). Výskyt adversních dat můžeme předpokládat na bázi vývoje podobného výskytu plevelů rezistentních k herbicidním zásahům. Problém primárně nastává se zásadním rozšířením technologie, kdy neobyčejně precizní selekce odolných rostlin umožňuje poměrně rychlé vyplnění vytlačených populací těmi, které disponují resistencí jakožto příkladnou evoluční výhodou (Owen & Zelaya 2005). Základem pro projev adversních dat tedy je namnožení morfortypů potenciálně nerozeznatelných algoritmem.

Srovnání vývoje odolnosti k obou technologiím může v současnosti působit irelevantně, ovšem při srovnání využití strojového vidění a totálních herbicidů v principu vychází velmi podobně. Adaptace pro plevelné rostliny v rámci strojového vidění může potenciálně brát formu vizuální mimiky plodin, čímž by algoritmus ztratil možnost spolehlivě oddělit obě kategorie. Vzhledem k rychlosti rozvoje resistencí se v případě herbicidů situace proměnila v praktický závod ve zbrojení mezi výrobci účinných látek a rostlinami (Owen & Zelaya 2005), jehož alternativa je u strojového vidění velmi náročná díky potřebě dostatečného množství dat zaznamenávajících probíhající změny. Dalším aspektem je vývoj ve vztahu k lokálním půdním a klimatickým podmínkám, které v zásadě umožňují fundamentální odlišnost stejných druhů v perspektivě algoritmu, čímž by se objem potřebných dat dále zvětšoval. Vzhledem k výše zmíněnému je v současnosti kompletní nahrazení plošných zásahů strojovým viděním v současnosti nepravděpodobné vzhledem absenci dostatku dat korespondujících konkrétním lokálním specifikům zajišťujícím spolehlivost modelu.

Empirická data poukazují na dlouhodobou obtíž při implementaci strojového učení, která vychází z obecných systematických nedostatků technologie. Samotný rámec adversních efektů v současnosti představuje dvojice hlavních obtíží. V první řadě se jedná o distribuční rozdíly mezi tréninkovými a cílovými daty, které narušují jejich předpokládanou uniformitu (Špelda & Střítecký 2021). Problémem zůstává současná závislost algoritmů na obdobné variabilitě proměnných, kdy předpokládané základní elementy (Ground truths) jsou v dynamickém prostředí nespolehlivou oporou. V druhé řadě jsou samotní adversní efekty zásadními narušiteli stability základních elementů statistické variability datových vstupů v situaci implementace modelu. Celkově se narušení stability či změny kovariance mohou vyskytovat společně a vést k dalšímu rozvoji disrupcí předpokládané uniformity dat (Špelda & Střítecký 2021). Konkluzivně je problematika systematicky natolik obsáhlá v jádru strojového učení, že pro budoucí implementaci lze počítat primárně s kombinací využití strojového vidění a dalších nástrojů, případně lidské kontroly procesu.

Přesto, že může otázka odolnosti modelů působit jako zásadní překážka pro využívání technologie v praxi, v realitě se stále jedná o potenciální řešení řady problémů. Rámec adversních efektů je primárně směřován k otázce dlouhodobé udržitelnosti technologického řešení vzhledem k současnému rozsahu agrární sféry. V první řadě je otázka budoucí udržitelnosti do určité míry mitigovatelná skrze rozvoj používání více spektrálních kanálů, které by do značné míry zajišťovali náročnější mimiku vzhledu plodin. V druhé řadě zůstává pro zlepšení odolnosti řešení možnost omezení implementace na prostorové zpřesnění plošných zásahů, kdy by riziko mimiky vzhledu plodin zůstávalo limitované vzhledem k selektivitě látek vůči plodinám. Logika druhé cesty by do značné míry čerpala inspiraci z formování fungicidních přípravků, které rezistenci brání kombinováním účinných látek a jejich aktivním mechanismem zasažení cílového organismu.

Obecně v oboru zůstává zásadním aktérem lidský operátor, jehož zařazení do současného procesu lze skrze strojové vidění pouze upravit, čímž by dosáhl proces stavu „human in the loop“ (Li et al. 2014). Navzdory reálné nedokonalosti lidské kontroly fungování technických řešení je zachování lidského nadhledu poměrně zásadním faktorem robustnosti vzhledem k odlišnému vnímání reality lidmi a stroji. Problematika dlouhodobé udržitelnosti je konstantně podrobena nejistotě budoucího vývoje, kdy je v současnosti výzkum adversních dat plnohodnotnou součástí oboru strojového učení (Madry et al. 2017). Lze tedy v delším časovém horizontu potenciálně očekávat tvorbu robustních či adversně rezistentních přístupů k strojovému vidění.

3.4 Směrování výzkumného rámce

Vzhledem k empirickým poznatkům získaným z rámcové literatury je v rámci analýzy možné předpokládat zaměření na širší portfolio lokálně významných rostlin v období výzkumu dle Dyrmann et al. (2016). Z hlediska zpracování a sběru dat práce míří k následování historicky ověřených metod při přípravě a anotaci snímků. Jasným předpokladem pro fungování algoritmu bude jeho stavba na základě umělé konvoluční neuronové sítě, doplněné o srovnávací model YOLOv3 jako ukázkou soudobého prakticky používaného nástroje (Redmon & Farhadi 2018). Celkově je směr výzkumu dle literatury veden s cílem maximalizace výkonu při detekci. Finálním výstupem práce by měl být komplexní přístup k problematice a fungující model analýzy vizuálních dat potenciálně aplikovatelný do praktického prostředí.

4 Metodika

Úspěšná tvorba algoritmu strojového učení stojí v principu na třech základních pilířích (Goodfellow et al. 2016). V první řadě je pro analýzu zásadní sběr dat, a to v kvalitativní i kvantitativní rovině. Nasbírané snímky je třeba následně podrobit ručnímu třídění a anotaci dle vybraných tříd objektů, tedy v případě analýzy rostlinných druhů napříč plevelným spektrem. Třetím a finálním krokem při tvorbě modelu je jeho trénování, při kterém je do modelové architektury neuronové sítě vložen dataset a jeho anotace. Následný průběh tréninku prezentuje modelu jednotlivé aspekty snímků, které procházejí opakovaně predeterminovanou sadou filtrů umožňující algoritmu zobecnění vybraných charakteristik objektů. Na základě procesu filtrování a selektivního rozboru snímků je následný algoritmus schopen detekovat vybrané objekty na validačních a testovacích součástech sesbíraných dat.

4.1 Sběr dat

Základním kamenem analýzy je zcela jednoznačně získávání relevantních a dostatečně objemných dat. V případě detekce objektů z vizuálních vstupů slouží, jakožto primární datová platforma určitá forma digitálního snímku, který může obsahovat kombinaci či výběr spektrálních pásem v závislosti na požadované vlastnosti (Liakos et al. 2018). V případě analýzy byla vybrána rovina infračerveného spektra (NIR), jejíž výběr byl primárně motivován snadnou eliminací necílových objektů a pozadí. Snímkování bylo v rámci výzkumu koncipováno jako pozemní platforma, která oproti dálkovému průzkumu umožňuje zřetelně kvalitnější zobrazení cílových objektů (Wang et al. 2019). V druhé řadě byl odklon práce od dálkového průzkumu motivován i relativně nižší komplexností proveditelnosti sběru dat, přičemž alternativou by byla i přidaná investice do pořízení a úpravy dronu.

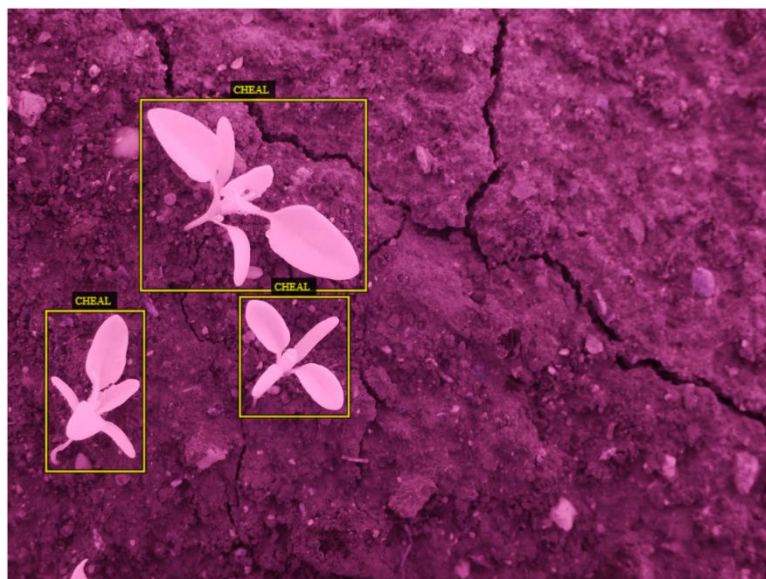
Snímkování samotné probíhalo především na experimentálních plochách ČZU v Praze Suchdole, disponujících přiměřenou diverzitou plevelové vegetace, což umožnilo získat dostatečného množství snímků pro jednotlivé kategorie. Celkově bylo do snímkování zahrnuto 18 druhů plevelných rostlin, které lze v kontextu polního hospodářství v České republice považovat za signifikantní pro ovlivnění výnosu plodin, zejména okopanin. Přestože je případné praktické zasazení výsledného modelu koncipováno pro samohybnou platformu, bylo v rámci sběru dat vzhledem k rozsahu práce využito primárně „přímého“ snímkování z rukou během pěšího přesunu skrze pozemek. Vzhledem k zájmu o minimální invazivnost sběru dat vůči sledovanému pozemku se pěší varianta projevila jakožto optimální přístup k využívaným plochám. Snímkování samotné bylo prováděno pomocí upraveného fotoaparátu Panasonic G5, který prošel odstraněním interního filtru, který blokuje infračervené záření a následně byl doplněn o vnější filtr Hoya R72 blokující viditelné záření do vlnové délky 720 nanometrů. Fotoaparát byl vybaven objektivem Panasonic G Vario 14 - 45 mm f/3,5-5,6 .

Průběh snímkování v polních podmínkách byl do značné míry ovlivněn kombinací vnějších faktorů, jejichž vliv je pro většinu vizuálních platforem častou slabinou (McCarthy et al. 2010). V případě infračerveného záření se problematika vnějšího prostředí projevila v kombinaci přílišného množství záření a odrazivosti půdy v pozadí. Pro získání kvalitativně použitelných snímků se sběr realizoval především v podmínkách částečně zavlhčené půdy, která zvýšila kontrast mezi rostlinami a pozadím a dále zatažené obloze, eliminující dopad přímého slunečního záření na snímané prostředí a vytváření stínů. Samotný sběr dat byl prováděn ve formě snímků vytvářených ze vzdálenosti 50 cm v kolmém směru vůči povrchu půdy při ohniskové vzdálenosti objektivu 45 mm. Rozlišení snímků bylo nastaveno na 3264×2248 pixelů a jejich datová velikost dosahovala přibližně 4 MB.

4.2 Příprava a anotace snímků

V druhé fázi zpracování nasbíraných dat bylo pro úspěšné trénování modelu zásadní dosažení dvou cílů. V první řadě se jednalo o pracovní rozdělení snímků na cílové objekty a pozadí, přičemž následně byly jednotlivé objekty anotovány, posléze doplněné o druhou fázi kvantifikace a rozdělení dat na tréninkový, validační a testovací dataset. V průběhu ruční selekce nasbíraných dat byly eliminovány neostré a jinak nekvalitní snímky, jejichž přítomnost by ohrožovala robustnost modelu (Dos Santos Ferreira et al. 2017), přičemž i ve vztahu k anotacím se jednalo o zjednodušení procesu. Následně byly sečteny jednotlivé druhové kategorie, ke kterým se směřovalo případné doplnění chybějícího počtu snímků s cílem dosažení nejlepšího možného poměru všech zastoupených tříd.

Po dosažení dostatečného množství zachycených objektů pro požadované třídy se realizovala jejich anotace, která byla prováděna pomocí „open source“ počítačového software VIA od Visual Geometry group. Jednotlivé anotace byly implementovány ve formě takzvaných anotačních boxů, tedy obdélníků či čtverců překrývajících hranami okrajové body označovaného objektu (viz Obrázek 6). Vzhledem k formě datasetu obsah anotací tvořily jak samostatné, tak kombinované nebo překrývané objekty, které byly označeny individuálními boxy. K označení sledovaných tříd (rostlinných druhů) bylo využito tzv. EPPO kódů pro zjednodušení formátu a snadnou aplikovatelnost díky přehlednosti systému. Jednotlivé anotace byly formátovány do souboru typu Json nebo alternativně YOLO, určených pro jednotlivé modely a zařazeny do samostatného adresáře.



Obrázek 6 Anotace objektů na snímku

Finálním krokem v přípravě dat pro trénování modelu byla multiplikace originálních dat a následné rozdělení pro jednotlivé operace. Znásobení celkového počtu snímků bylo provedeno skrze metodu duplikace při rotaci os, kdy byly náhodně vybrané snímky upraveny pootočením o devadesát stupňů. Skrze výše zmíněnou operaci se pro účely modelu finální dataset efektivně zdvojnásobil, čímž podpořil komplexní robustnost výsledného algoritmu (Farooq et al. 2019). Rozdělení datasetu na tréninkovou, evaluační a testovací část bylo dle logiky maximalizace úspěšnosti detekce orientováno na zásadní převahu tréninkových dat (Miikkulainen et al. 2019). Z celkového počtu bylo tedy náhodně vybráno 70 % anotovaných snímků, které byly přiřazeny pro trénování. Dalších 10 % bylo obdobným způsobem vybráno pro evaluaci modelu a zbylých 20 % snímků sloužilo pro testování výsledné úspěšnosti v detekci a rozpoznání jednotlivých objektů.

4.3 Tvorba modelů a jejich trénink

Proces tréninku modelů byl v zásadě diktován dvojitou logikou. Za prvé se řídil snahou o dosažení co nejvyšší úspěšnosti v detekci objektů a za druhé cílem komparativního ověření výkonnosti vlastního modelu neuronové sítě. Druhého cíle bylo dosaženo pomocí tréninku třetí verze výkonného modelu You only look once (YOLOv3) na stejných datech s identickým rozdělením jednotlivých sub-datasetů. Rámcově oba procesy probíhaly v jednotném prostředí, čímž byla zajištěna jednotnost testovacích podmínek, kdy jediným zásadním rozdílem při tréninku obou modelů byl přístup k změně rozměrů tréninkových dat. V případě vlastního modelu bylo použito modulu PIL v sadě nástrojů jazyka Python, oproti kterému model YOLOv3 provádí změnu rozměrů automaticky skrze svou software bázi Darknet.

Trénink obou využitých modelů se realizoval na částečně upraveném herním počítači, který byl pro účely analýzy formátován a vybaven operačním systémem Ubuntu, který je distribucí systému Linux. Z hardwarového hlediska byl počítač vybaven procesorem Intel i7, 32 GB RAM a grafické karty NVIDIA GeForce GTX 1080 s vnitřní architekturou „Pascal“. Pro oba modely bylo využito programovacího jazyka Python 3.8 s knihovnami Keras 2.0, Matplotlib 3.5.1 a Numpy 1.18.0. Trénink samotný byl v obou případech prováděn za užití výkonu grafické karty (GPU) pomocí nástroje Tensorflow 2.8 z dílny týmu Google brain, umožňující rychlejší a efektivnější zpracování prezentovaných dat (Abadi et al. 2016). Finálním aspektem fungování Tensorflow bylo pro práci připojení nástrojů CUDA 6.1 a knihovny CUDNN 8.3, umožňující plnohodnotné využití potenciálu grafické karty pro trénink.

4.3.1 Vlastní model

Ve srovnání s konvenčními prostředky typu SVM, můžeme hluboké neuronové sítě postavené na konvolučních vrstvách považovat za mnohonásobně výkonnější alternativu. Pro řešení práce bylo tedy vybráno odpovídající architektury, schopné dosažení robustních výsledků v sledované skupině dat. Rámcově byla vlastní architektura částečně inspirována modelem užitým při detekci širokolístých plevelů v pastvinách (Zhang et al. 2018) a modelem formulovaným pro detekci plevelů v prostředí Dánska (Dyrmann et al. 2016). Ve srovnání s inspirací ovšem architektura modelu užitého v práci využívá odlišný počet konvolučních vrstev a vyšší koncentraci filtrů na jednu vrstvu. Rámcově se vlastní architektura práce od vzorových projektů liší i v rovině jednotlivých vrstev, kdy struktura finální, výstupové vrstvy byla postavena odlišně od obou projektů (viz kód).

```

def main_model():

    main_input = Input(shape=(416,416,1))
    L1 = Conv2D(32, (6,6), activation='relu')(main_input)
    L2 = Conv2D(64, (6, 6), activation='relu')(L1)
    L3 = MaxPool2D()(L2)
    L4 = BatchNormalization()(L3)

    L5 = Conv2D(128, (3,3), activation='relu')(L4)
    L6 = Conv2D(256, (3,3), activation='relu')(L5)
    L7 = MaxPool2D()(L6)
    L8 = BatchNormalization()(L7)

    L9 = Conv2D(32, (3, 3), activation='relu')(L8)
    L10 = Conv2D(64, (3, 3), activation='relu')(L9)
    L11 = MaxPool2D()(L10)
    L12 = BatchNormalization()(L11)

    L13 = GlobalAvgPool2D()(L12)
    L14 = Dense(64, activation='relu')(L13)
    FL = Dense(19, activation='softmax')(L14)

    model = tensorflow.keras.Model(inputs=main_input, outputs=FL)

    return model

```

Ústředním elementem modelu byla v rámci obrazové analýzy dvourozměrných snímků konvoluční vrstva, která vstupní data vystavuje predeterminovanému množství filtrů o selektovaných rozměrech. V rámci hyperparametrů došlo v modelu k nárůstu počtu filtrů s postupem konvolučních vrstev (viz kód), doplněného o částečný pokles velikosti segmentů s cílem zvýšení přesnosti analýzy. Jednotlivé konvoluční vrstvy obsahovaly vedle samotných filtrů i aktivační funkci, jejíž cílem bylo narušení linearity (Farooq et al 2019) . Pro vlastní architekturu byla vybrána aktivační funkce ReLU (rektifikovaná lineární jednotka), aplikující nelineární funkci na výstup přiřazené konvoluční vrstvy. Výběr funkce byl motivován především zefektivněním výpočetní rychlosti za cenu minimálního poklesu přesnosti, kdy většina alternativ jednoznačně zvyšuje zatížení výpočetní techniky (He et al. 2018).

Jednotlivé konvoluční vrstvy v architektuře doplňovaly vrstvy maximalizace výstupů (Maxpool), jejichž výstupní data obsahují pouze nejvyšší hodnotu ze sledovaných vstupních parametrů (Dyrmann et al. 2016). Vrstvy maximalizace výstupů obsahovaly filtr o rozměrech 2x2, který v průběhu experimentálních tréninků modelu dosahoval nejlepších výsledků. Pro efektivní udržení konstantního rozsahu vstupů do jednotlivých vrstev, tak aby nebyly ovlivněny aktualizací předešlých výstupů, model obsahoval vrstvy normalizace dávek (batch normalization), provedená pomocí porovnání výstupu vrstvy s její směrodatnou odchylkou (Ioffe & Szegedy 2015). Výsledná forma dat odcházejících z vrstvy pomohla zredukovat počet nutných tréninkových iterací, čímž dále zvýšila efektivitu modelu bez snížení jeho přesnosti.

Závěrečnou vrstvou neuronové sítě tvořila trojice vrstev tvořená jednou globálně průměrující vrstvou a dvojicí „zahušťovacích“ vrstev (viz kód). Globální průměrování (Global Average pool) posloužil modelu k zprůměrování výstupních dat z předešlého segmentu a jejich nasměrování ke dvojici výstupních vrstev. Globální průměrování bylo vybráno namísto dříve používané metody plně propojených vrstev s cílem dosažení vyšší robustnosti modelu a jeho přirozenější koordinaci s konvolučními vrstvami (Hsiao et al. 2019). Závěrečná dvojice „zahušťovacích“ vrstev (Dense layer) dokončila práci předcházejících neuronů formulací výsledných kategorií a přiřazením datových výstupů do vektorů odpovídajících kategorií. Druhá a závěrečná z těchto vrstev již primárně navracela výsledky zpět k výstupu funkce.

Celkově byl koncipovaný model trénován po období 150 epoch, během kterých jednotlivé vrstvy postupně detekovaly menší segmenty objektů, ze kterých byl postupem času složen kompletní tvar. Úvodní filtry tedy sloužily primárně k odhalení hran pomocí kontrastních pixelů, přičemž se s každou normalizací dávek posunula rovina detekce k liniím a následně celkovým tvarům (Zhang et al. 2018). V průběhu tréninku se celkový počet nastudovatelných parametrů vyšplhal přes hranici jednoho milionu, což ovšem obdobně jako v analýze Dyrmann et al. (2016) zdaleka nedosahovalo počtu parametrů sledovaných při tréninku Alexnetu.

4.3.2 YOLOv3

Přestože se v případě YOLOv3 jednalo o použití předpřipraveného nástroje bylo pro úspěšné použití nutné virtuální prostředí řádně připravit. Proces implementace modelu na sdílený dataset vyžadoval několik základních kroků, nutných pro úspěšnou realizaci tréninku. V první řadě bylo nutné převést formát anotací jednotlivých snímků do optimální formy pro YOLOv3, přičemž bylo využito opětovně nástrojů software VIA. Následná optimalizace modelu probíhala v prostředí Microsoft Visual Studio 17.0, které posloužilo jako báze pro celý trénink modelu. Zásadní zrychlení modelu zajistila jeho předtrénovaná verze přístupná z repositáře autorů, díky níž výsledný algoritmus disponoval vyšší robustností doplněnou o zrychlení tréninkového procesu (Wang et al. 2019). Trénink samotný byl spouštěn pomocí krátkého skriptu, přičemž bylo pro jeho úspěšný průběh potřebné zajistit stabilitu operačního systému a správné umístění tréninkového a evaluačního datasetu. Nastavení tréninku neuronové sítě se řídilo výsledky vlastního modelu, který diktoval omezení v rámci rozsahu epoch.

Po úspěšné realizaci tréninku byla provedena evaluace a testování modelu na identických sadách datasetů stvrzující výkonnost systému a prezentující výsledky v reálném čase. Samotná proces evaluace byl realizován v rámci prostředí tréninkového nastavení modelu, kdy Darknet-53 obsahuje automatický evaluační nástroj (Redmon & Farhadi 2018). Souběžnost procesu evaluace s tréninkem vyústila v jednotný výstupní soubor dat, které byly následně ověřeny skrze vstup testovacích dat. Testovací nástroj v předepsaném režimu vložil daná data do neuronové sítě a následně vypsál tabulku obsahující výsledné hodnoty přesnosti a spolehlivosti algoritmu. Rámcově byly jednotlivé kroky realizovány uvnitř prostředí Darknet-53, které při relativně snadné skriptové implementaci nabízí nástroje pro všechny potřebné procesy.

4.3.3 Statistické zhodnocení úspěšnosti

Pro evaluaci výstupů modelů bylo použito několik základních statistických parametrů. V první řadě se jednalo o praktické posouzení úspěšnosti detekcí, kdy přímou přesnost výsledků tvořil posudek detekcí odvozený z výskytu chyb I. a II. typu či záměny tříd. Chyba I. typu pro algoritmus znamenala přisouzení detekce do prostoru, kde žádný sledovaný objekt nenachází. Chyba II. typu naopak popisovala stav, kdy model úspěšnou detekci navzdory výskytu objektu neprovedl. Záměna tříd pro model znamenala stav, kdy navzdory úspěšné detekci objektu došlo k nesprávnému přiřazení tříd, a tedy k nesprávné identifikaci sledované rostliny. V druhé řadě byly modely v rámci evaluace měřeny i výkonnostně pomocí srovnání času nutného k zpracování vstupních dat a poskytnutí výsledků.

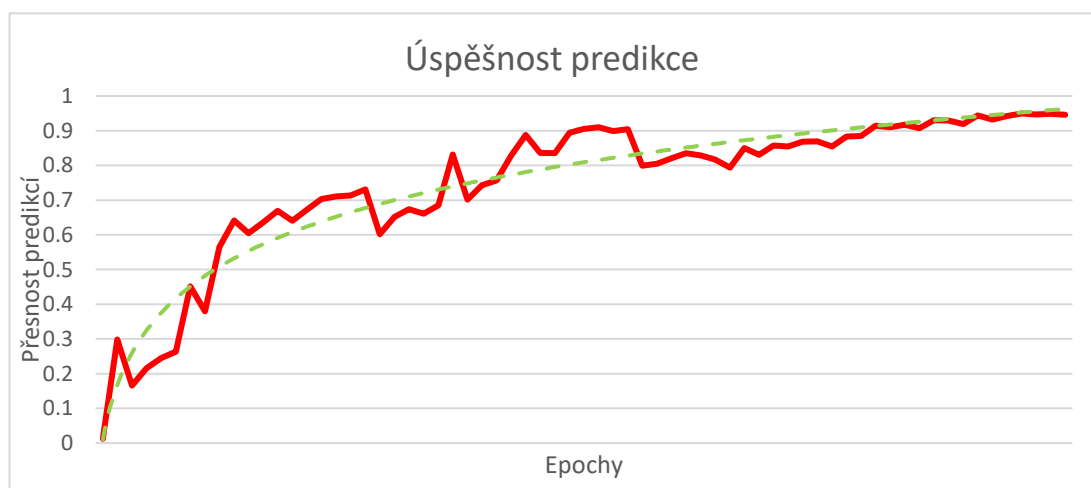
5 Výsledky

Kapitola je rozdělena na dva primární bloky reflektující výsledky dvou aplikovaných algoritmů a jejich jednotlivé elementy. První z podkapitol rozebírá výsledky implementace vlastního modelu v relevanci k detekci objektů a správného přiřazení jednotlivých tříd. Výsledná přesnost je v podkapitole dále kontextualizována v relevanci k vizuálnímu vztahu jednotlivých sledovaných druhů a doplněna o obecný rozbor a výčet chybových výstupů modelu. Druhá podkapitola popisuje obdobným způsobem implementaci modelu YOLOv3 a jeho elementů, přičemž částečně reflektuje odlišnost přístupu obou modelů k zpracování datasetu.

5.1 Výsledky vlastního modelu

Rámcové výsledky vlastního modelu práce reflektují rozsah tréninku, datasetu a limity konvolučních vrstev. Průběh tréninku byl zastaven vzhledem k postupnému zploštění křivky validace výstupů (validation loss) a praktické maximalizaci rozdílu mezi validační a tréninkovou ztrátou (Dyrmann et al 2016). Realisticky lze dosáhnout plného zploštění křivky, ovšem ve vztahu k časovým a výkonnostním limitům je maximalistický přístup k tréninku neefektivní díky relativně nízkému nárůstu úspěšnosti na úkor nepoměrně vyšší časové a zdrojové investice. K bodu, kdy již nárůst hodnot přesnosti byl vzhledem k času zanedbatelný došlo v epoše číslo 150, která se stala i zdrojem procentuálně nejúspěšnějších detekčních vah.

Poslední epocha dosáhla v rámci evaluačního datasetu úspěšnosti klasifikace 94,96 %, která byla následně testováním na připraveném zbytku dat ověřena na hranici 94,5 %. Nárůst přesnosti v průběhu jednotlivých epoch následoval standardní trend (viz Graf 1.), na kterém je poměrně dobře znatelné zplošťování křivky úspěšnosti. V rovině testovacích výsledků se poměrně zajímavým způsobem projevil výskyt chyby I. Typu, která se zhruba v 45 % neúspěšných detekcí objevila díky záměně posklizňových zbytků s plevelnými rostlinami. Celkově se v rámci tréninkového vývoje podařilo na evaluačních datech získat poměrně vysokou přesnost detekce objektů. V rovině výkonnosti byl ovšem vlastní model mírně pomalejší oproti očekávání, kdy zpracování snímku zabíralo 28 ms.



Graf 1 Záznam přesnosti modelu na tréninkových datech

Z hlediska jednotlivých druhů se pro model jako nejsnáze rozpoznatelný projevil druh *Lamium amplexicaule*, který se modelu podařilo správně detekovat v 99,2 % případů. Na druhém a třetím místě v obdobné úrovni přesnosti se umístily *Chenopodium album* s 98,9 % a *Fallopia convolvulus* s 98,8 %. Na opačném konci žebříčku v chybovosti dominoval *Thlaspi arvense*, jehož správné určení se modelu podařilo pouze v 65,2 % případů. Obdobných výsledků dosáhl model i v případě *Tripleurospermum inodorum* a *Fumaria officinalis*, kde se úspěšnost pohybovala na úrovni 80,5 a 84,8 %. Za poměrně specifickým výsledkem v rámci modelu stál i případ *Echinochloa crus-galli*. Na jedné straně dařilo modelu poměrně přesně detekovat jedince s úspěšností 89 %, ovšem na straně druhé byl druh zodpovědný za prakticky veškeré případy výskytu chyby I. typu u evaluačního i testovacího datasetu.

5.2 Výsledky modelu YOLOv3

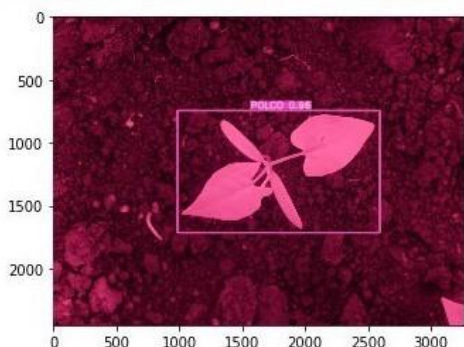
V případě modelu YOLOv3 byl obdobně jako o vlastního modelu pro trénink stanoven rozsah 150 tréninkových epoch. Srovnatelný rozsah byl vybrán s cílem vyrovnanosti podmínek modelů navzdory reálnému potenciálu YOLOv3 k pozdějšímu zploštění křivky. Průběh vývoje validační křivky byl vůči vlastnímu modelu relativně obdobný s mírnou variací vzniklou díky regresní logice, na níž je celý model vybudován (Redmon et al. 2016). Trénink modelu se výsledně projevil svou rychlostí již v průběhu tréninku, kdy období 150 epoch zabralo přibližně o 20 % času méně než u vlastního modelu. Obdobně jako v případě vlastního modelu byla výsledkem tréninku sada „nejlepších“ detekčních vah, jejichž hodnoty odpovídaly výsledku poslední z epoch.

| Epoch | gpu_mem | box | obj | cls | labels | img_size | | |
|---------|---------|---------|---------|---------|--------|-----------|-------|--|
| 150/150 | 0.946G | 0.01539 | 0.01026 | 0.00411 | 55 | 320: 100% | █ | |
| | Class | Images | Labels | P | R | mAP@.5 | mAP@ | |
| | all | 1411 | 1806 | 0.984 | 0.979 | 0.88 | 0.782 | |

Obrázek 7 Hodnoty 150 epochy tréninku modelu

Úspěšnost detekce poslední tréninkové epochy modelu byla na validačním datasetu rovna hodnotě 98,45 % (viz Obrázek 7). Vývoj jednotlivých epoch následoval poměrně lineární trend nárůstu, k jehož narušení došlo pouze v okolí 40 epochy, kdy přesnost klesla o přibližných 10 %. Finální úroveň přesnosti byla ověřena na testovacím datasetu, kdy model dosáhl úrovně 97,92 % úspěšnosti. V průběhu procesu testování se obdobně jako u vlastního modelu objevily případy chyby I. typu, ovšem poměr jejich výskytu byl nižší na přibližně 15 % případů. Z hlediska rychlosti byl model YOLOv3 skutečně velmi výkonný, kdy dosahoval procesního času 22,8 ms na jeden snímek (viz Obrázek 8).

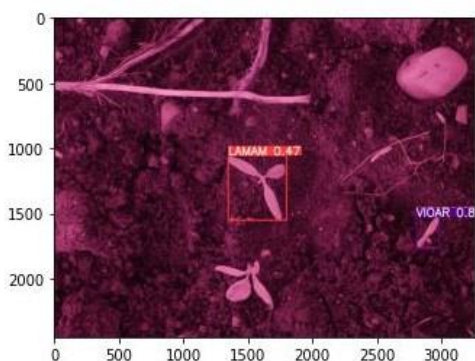
image 1/1: 2448x3264 1 POLCO
Speed: 131.1ms pre-process, 22.8ms inference, 1.3ms NMS per image at shape (1, 3, 480, 640)



Obrázek 8 Výstup modelu na snímek Fallopia convolvulus

Model jednotlivé druhy sledovaných rostlin detekoval s relativně nízkou variabilitou ve většině případů. Nejpřesněji detekované byly pro model snímky obsahující *Fallopia convolvulus*, kde průměrná úspěšnost přesahovala 99,5 %. Hranici 99 % úspěšnosti detekce model znatelně překročil i v případě snímků *Chenopodium album*. Většina zbylých tříd se u modelu pohybovala v rozmezí 98 a 99 % úspěšnosti detekce, kdy výjimku tvořila pouze kombinace *Papaver rhoeas* a *Thlaspi arvense*. U obou druhů docházelo v testovacím datasetu k poměrně časté záměně a chybné identifikaci, která vedla k celkové úspěšnosti pro oba druhy pod 90 %. Vedle výskytu chyby I. typu a II. typu se u testovacího datasetu vyskytovala i druhová záměna, kdy model přiřadil objekt do nesprávné třídy (viz Obrázek 9), která navzdory menšímu počtu případů upozorňuje na limity tréninkového datasetu.

image 1/1: 2448x3264 1 VIOAR, 2 LAMAMs
Speed: 133.3ms pre-process, 22.7ms inference, 1.3ms NMS per image at shape (1, 3, 480, 640)



Obrázek 9 Výsledek modelu s chybou II. typu a chybným přiřazením třídy

6 Diskuze

Výsledky obou modelů je v první řadě nutné porovnat v rovině jejich technických vlastností, robustnosti a celkové přesnosti v detekci. Srovnání pomůže ověřit nejen použitelnost vstupních dat, ale i konkurenceschopnost vlastního modelu. Výsledky lze dále porovnávat s obdobnými projekty zaměřenými na problematiku detekce plevelů pomocí strojového vidění, čímž se zabývá druhá podkapitola. Rámcové srovnání je možné následně částečně použít jako bázi pro finální rozbor pro implementační potenciál modelového řešení, který je hlavním elementem třetí podkapitoly. Samotná třetí podkapitola se následně dělí dle roviny analýzy na ekonomicky formulovaný rozbor empirických dat a teoreticko-praktickou posouzení robustnosti technologie.

6.1 Komparace modelů

Primárním cílem tréninku modelu YOLOv3 na identických datech bylo ověření jejich použitelnosti pro strojové učení. Druhým a doplňujícím cílem sloužilo ověření výkonnosti vlastního modelu v konkurenci soudobého prakticky používaného modelu. Bezpochyby lze protiargumentovat slabými stránkami YOLOv3, mezi které patří například snížená robustnost ve vztahu k přesné detekci většího počtu menších objektů ve skupině či přímá identifikace soliterních objektů (Redmon et al. 2016). Celkově se v kontextu rozhlasu práce podařilo vytrénovat pouze jeden alternativní model, který je ovšem principiálně dostačující. Konkluzivně tedy byl model vybrán především pro jeho rychlost, která umožňuje minimálně v jedné rovině srovnání s přibližným vrcholem výkonnosti.

Základním srovnatelným měřítkem obou modelů zůstává jejich architektura a její rozsah. V přímém srovnání je zcela jednoznačný rozdíl v rozsahu, kdy 53 konvolučních vrstev Darknet-53 zcela jednoznačně převyšuje rozsah vlastního modelu. Ve smyslu architektury lze dále vyzdvihnout posun v užití plně propojených vrstev u YOLOv3, které vlastní model nahrazuje pokročilejší kombinací globálního zprůměrování a zahušťovacích vrstev. Rámcově tedy vlastní model působí při srovnání s YOLOv3 jako lehčí a pružnější model, čímž by bylo možné vysvětlit jeho výsledky v celkovém porovnání. Úvodním elementem, který přímo vychází z architektury je rychlost v procesování dat, kde zcela jednoznačně vítězí YOLOv3, neboť se jedná rychlostně vrcholný soudobý model. Při srovnání 22,7 ms s 28 ms vlastního modelu je výsledek poměrně jednoznačný, ovšem v celkovém kontextu lze rozdíl 5 ms považovat za poměrně rychlostně výkonný.

Při srovnání přesnosti modelů je třeba se v první řadě zaměřit na fundamentální otázku přesnosti detekce cílových objektů. Přesto, že v absolutním měřítku výsledný rozdíl 3,32 % není nijak závratný, v hledisku přesnosti výstupů strojového učení se jedná o poměrně markantní rozdíl (Wang et al. 2019). Za zásadní hranici v přesnosti lze považovat obdobně jako u statistiky hranici 95 % (Nawara 2010), kterou vlastní model bohužel nepřekonal, ovšem které se velmi přiblížil. Model YOLOv3 naopak kýžené hranice dosáhl s relativně velkou rezervou. Celkově ovšem již přesázení hranice 90 % úspěšnosti lze považovat za úspěch vzhledem k prokázání stability modelu pro další rozvoj a zlepšení výsledků.

Poměrně specifickým elementem srovnání je segmentovaná úspěšnost identifikace jednotlivých tříd sledovaných objektů. Ačkoliv je přesný důvod rozdílných výsledků u jednotlivých prvků datasetu jen velmi složitě dohledatelný, je možné v rámci analýzy argumentovat primárně rozdílností použitých filtrů. Podobnost, kdy *Fallopia convolvulus* a *Thlaspi arvense* tvořily u obou modelů praktické protipóly v úspěšnosti detekce je vysvětlitelná dvojicí prvků. V první řadě se jedná o nepoměr v objemu tréninkových dat, kdy jedinci *Thlaspi arvense* tvořily ve srovnání se zbytkem tříd relativní minoritu. Naopak v řadě druhé lze polemizovat o obecné vhodnosti tvarových prvků obou druhů, kdy zcela jednoznačně existují lépe a hůře identifikovatelné tvary pro strojové vidění (Dyrmann et al. 2016). Značnou variabilitu přesnosti klasifikace tříd zbylých druhů mezi oběma modely lze primárně vysvětlit v rovině výše zmíněných filtrů. Detekce některých druhů v pixelovém měřítku lépe fungovala s menšími filtry použitými v YOLOv3, u jiných naopak lépe vyhovovala kombinace větších filtrů vlastního modelu.

Finální srovnávacím elementem obou modelů by měla být jejich odolnost k chybám. Ačkoliv výsledky testování na vybrané části dat nezaznamenalo zásadnější míru adversních efektů narušujících stabilitu výsledků, je nezbytné analyzovat i selekci chybných výsledků. Při přímém srovnání je zcela jednoznačně znatelný nepoměr výskytu chyby I. typu u vlastního modelu oproti relativně rovnoměrnému rozložení příčin chybných výsledků u modelu YOLOv3. Z hlediska příčin zde opětovně jako nejpravděpodobnější působí odlišná velikost filtrů konvolučních vrstev, kdy *Echinochloa crus-galli* může poměrně snadno splynout v širším měřítku s posklizňovými zbytky. V rovině chyby II. typu a záměně druhů byla u obou modelů v majoritě příkladů prezentovaná jistota detekce poměrně nízká, což lze prakticky eliminovat zvýšením minimální hodnoty jistoty modelu pro zobrazení zaznamenaných detekcí.

6.2 Srovnání s obdobnými modely

V současnosti existuje poměrně široká skupina obdobně směřovaných výzkumů, jejich podobnost ovšem naráží při pokusu o porovnání na fundamentální rovinu odlišností. Zásadní rozdíly vedoucí k odlišnosti podobných analýz se skládají z kombinace odlišných forem dat, druhového spektra či formátu snímků, jejichž kombinace mají v zásadě minimum průsečíků. Přesto, že přímé srovnání jednotlivých modelů a jejich výsledků je značně problematické díky řadě faktorů, jeho provedení je pro posouzení splnění cílů výzkumu nutné. Pro tedy zůstalo klíčovým kritériem zaměření na detekci plevelných rostlin v polních podmínkách za pomoci digitálních snímků.

Jako bázi srovnání lze vzít 95 % úspěšnost v analýze Alam et al. (2020), která ve srovnání s vlastním modelem pracuje se snímky ve formátu RGB. Přístup autorů k eliminaci pozadí ovšem staví na obdobné logice jako vlastní analýza, je ovšem možné argumentovat o jeho nižší efektivitě ve srovnání s infračervenými snímky. Samotný model je ve srovnatelně výkonný v rovině přesnosti, ovšem jeho tréninková báze je zcela jednoznačně limitovanější s 97 snímky na každou z tříd (Alam et al. 2020). Rychlostně je vzhledem ke své kompaktnosti model relativně srovnatelně rychlý, kdy dochází k rychlejší segmentaci vzhledem k odlišnosti architektury. Druhové spektrum je opětovně vzhledem k lokalizaci výzkumného týmu naprosto odlišné, eliminující přímé srovnání na bázi jednotlivých tříd.

Dalším potenciálně srovnatelným příkladem může být výsledná přesnost datasetu B 96 % v analýze Farooq et al. (2019). Analýza využívá data v hyperspektrálním formátu, přičemž jádro detekce je zkombinované z tvarových a barevných prvků jednotlivých objektů. Výhodou formátu je oproti infračervenému spektru použitým v analýze kombinace relativně snadná eliminace pozadí a vyšší robustnost detekce díky prvkům barevného spektra. Architektura obou modelů je relativně srovnatelná, ovšem její výkonnost mimo přesnost detekcí lze srovnávat pouze velmi těžko vzhledem k absenci údajů v článku popisujícím experiment. Srovnání dat komplikuje v první řadě opětovná odlišnost druhového spektra, ovšem v druhé i dvojice použitých datasetů. Rozsah obou datasetů, podobně jako počet tříd je limitovaný, čímž lze následně argumentovat v komparaci výsledné přesnosti modelů.

Jako finální srovnání pro výsledky vlastního modelu poslouží 86,2 % přesnost modelu dle Dyrmann et al. (2016), jehož druhový rozsah i lokalizace jsou z trojice srovnávaných modelů výslednému nejbližší. Zásadní odlišností zůstává využití RGB dat, která zůstávají samostatně více náchylná k záměně pozadí se sledovanými objekty. Rozsah datasetu přesahující po úpravách objem 10000 objektů ve 22 třídách lze ve srovnání s prezentovaným objemem dat považovat za odpovídající velikosti projektu, kdy rozsah diplomové práce má vlastní limity. Vývojový posun v technologii se reflektuje i v rovině obou modelů na rozdílech v jejich celkové výkonnosti, kdy prezentovaná přesnost odpovídá rozvoji technologie strojového učení. Finálně srovnání modelů podporuje logiku technologického a metodického posunu, který lze reflektovat rozvojem datových platform v podobě kombinace více spektrálních pásem či užití stereokamer a dalších přístupů.

Celkovým výstupem srovnání vlastního modelu v kontextu obdobných výzkumů může v první řadě být stvrzení jeho konkurenceschopnosti, kdy se velmi pravděpodobně nejedná o nefunkční přístup k problematice. V druhé řadě porovnání podporuje úvahu o využitelnosti infračerveného spektra pro snadnou eliminaci pozadí pro algoritmus, kdy užití čistých RGB dat podléhá riziku záměny půdy s objekty. Rovina rozsahu datasetu poukazuje při srovnání s výzkumy na problematiku získávání dostatečného množství referenčních dat, kde ovšem práce patří mezi lépe vybavené, a tedy i rámcově robustnější. Celkově je ovšem primární pro budoucí výzkum motivovat k dalšímu rozvoji rozsáhlého sběru a přípravy dat nutné pro úspěšnost a odolnost výsledných modelů. Z obecného hlediska kontinuální rozvoj výzkumu v oboru strojové detekce plevelných rostlin stvrzuje logiku výzkumu, který zcela jednoznačně zůstává v současné snaze o ekologizaci oboru smysluplný.

7 Závěr

Navzdory řadě problematických faktorů je možné konkluzivně potvrdit praktický úspěch práce v rámci jejích dílčích i primárních cílů. Realita implementování technologie v širokém měřítku do podmínek praxe, tedy v rámci škálování současné produkce zůstává v současnosti stále relativně limitovaná. Celkově je ovšem dlouhodobý trend v oboru důkazem reálného potenciálu technologie strojového vidění k fundamentálnímu vylepšení soudobého precizního zemědělství (Sharma et al. 2020). Vzhledem k výstupům práce je ovšem i samotné precizní zemědělství v současnosti značně podřízeno společenskému tlaku směřujícímu agrární sféru k uplatnění konceptu udržitelnosti v širším měřítku. Finálním aspektem výše zmíněného tlaku je i existující rámec finanční motivace, která ač částečně řízena společenskými aspekty dlouhodobě podléhá geopolitickým a tržním vlivům.

Z hlediska dílčích cílů práce úspěšně implementovala systém pro sběr snímků, které následně díky predeterminovaným procesům přípravy a anotace posloužily jako báze pro trénink dvojice modelů. Modelů samotných bylo užito k ověření použitelnosti infračerveného spektra pro detekci plevelů, přičemž jeho výhody a slabé stránky následně empiricky analyzovala kapitola diskuse. Oba modely souběžně splnily i primární cíl práce, tedy ověření možností morfologických znaků jednotlivých druhů rostlin jakožto rozlišovacích znaků pro automatické identifikace. Finálním cílem práce bylo v kontextu užití dvojice modelů ověření schopností vlastního modelu při přímém porovnání s prakticky implementovaným nástrojem. Komparace byla později doplněna o empirické srovnání s kombinací srovnatelně směřovaných výzkumů v oboru.

Sekundárním cíle diskuze byla i empirická analýza implementačního potenciálu technologie, jejíž výstupem byl rozbor ekonomický a technický. Z ekonomického hlediska se navzdory řadě komplikací v celkovém kontextu technologie strojového vidění jeví jakožto progresivně realistická v blízkém časovém rámci. Přesto, že se z hlediska politických motivací jedná o prioritní oblast, tržní prostředí neumožňuje revoluční uplatnění technologií vzhledem k snaze o maximalizaci zisků pocházejících ze soudobých přístupů. Na druhé straně technické hledisko nekontrolovaného prostředí s adversními efekty nabízí poměrně výrazné komplikace pro realitu aplikace technologie pro praktické nasazení. V kontextu současného stavu strojového učení a nepředvídatelnosti chování otevřeného systému agrocenózy není možné zaručit stabilitu výstupů technologie v delším časovém horizontu.

Z hlediska hypotéz práce došla k závěru:

H₁: Při využití neuronových sítí je možné detekovat plevelné a kulturní rostliny s jistotou odlišení přesahující 85 %.

Model je při použití soudobých nástrojů schopen hranici zcela jednoznačně překonat, kdy, jak vlastní, tak srovnávací model přesáhly hranici 90 % úspěšnosti. V kontextu praktického užití zůstává používání infračerveného spektra otázkou, vzhledem ke kombinaci jeho slabin, komplikujících možnost nelimitované praktické implementace. Celkově model v testovacích podmínkách dosahuje hodnot plně potvrzujících hypotézu.

H₂: Algoritmus strojového učení je schopen vzájemně rozlišit vybrané plevelné druhy s přesností přesahující 85 %.

Obdobně jako v případě první hypotézy byl výsledný produkt obou modelů schopen hranici zcela jednoznačně překonat, kdy problematika detekce dosahovala prakticky identických hodnot. Identifikace v širším kontextu do určité míry omezoval výskyt chyb, které ve vybraných případech s nízkou jistotou zaměnily některé druhy. Z hlediska chyby I. typu se u vlastního modelu projevil komparativně významný výskyt u *Echinochloa crus-galli*. Celkově ovšem oba modely dosahovaly uspokojivých hodnot, čímž základní hypotézu potvrdily.

Finálně je třeba v kontextu empiricky ověřených aspektů na výsledky analýzy pohlížet v širším měřítku. Navzdory relativně slibné úrovni přesnosti identifikace zůstává algoritmus limitován rozsahem dat, přičemž zásadním aspektem praktické aplikace je prostředí otevřeného systému. Pro úspěšné zařazení technologie do produkční reality agrární sféry by model musel disponovat rozsáhlejší druhovou i záznamovou bází a robustním zdrojem vizuálních dat. Přesto, že výstup analýzy není možné aplikovat do praxe ani v limitovaném měřítku, výstup v kontextu lokalizovaných podmínek a specifické volby senzorů zůstává krokem správným směrem. Na výsledky by mohl v budoucnosti navazovat další výzkum zaměřený na zakomponování silných elementů aplikovaného přístupu do širšího kontextu robustní a smysluplné technologie na kontrolu zaplevelení porostů.

8 Literatura

- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, ..., Zheng X. 2016. {TensorFlow}: A System for {Large-Scale} Machine Learning. Pages 265-283 in 12th USENIX symposium on operating systems design and implementation. OSDI 16.
- Abouzahir S, Sadik M, Sabir E. 2022. Lightweight Computer Vision System for Automated Weed Mapping. Pages 372-376 in 2022 12th Annual Computing and Communication Workshop and Conference (CCWC). IEEE.
- Ahmed F, Al-Mamun H A, Bari A H, Hossain E, Kwan P. 2012. Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection* **40**:98-104.
- Alam M, Alam M S, Roman M, Tufail M, Khan M U, Khan M T. 2020. Real-time machine-learning based crop/weed detection and classification for variable-rate spraying in precision agriculture. Pages 273-280 in 2020 7th International Conference on Electrical and Electronics Engineering (ICEEE). IEEE.
- Antoniuk L, Gernego I, Dyba V, Polishchuk Y, Sybirianska Y. 2017. Barriers and opportunities for hi-tech innovative small and medium enterprises development in the 4th industrial revolution era. *Problems and Perspectives in Management* **15**:100-113.
- Camargo E R, Zapiola M L, de Avila L A, Garcia M A, Plaza G, Gazziero D, Hoyos V. 2020. Current situation regarding herbicide regulation and public perception in South America. *Weed Science* **68**:232-239.
- Chen L, Zhang J G, Su H F, Guo W. 2010. Weed identification method based on probabilistic neural network in the corn seedlings field. Pages 1528-1531 in 2010 international conference on machine learning and cybernetics. IEEE.
- Davis G, Casady W W, Massey R E. 1998. *Precision Agriculture: An Introduction*. University of Missouri.
- dos Santos Ferreira A, Freitas D M, da Silva G G, Pistori H, Folhes M T. 2017. Weed detection in soybean crops using ConvNets. *Computers and Electronics in Agriculture* **143**:314-324.
- Dyrmann M, Karstoft H, Midtiby H S. 2016. Plant species classification using deep convolutional neural network. *Biosystems engineering* **151**:72-80.
- Farooq A, Jia X, Hu J, Zhou J. (2019). Multi-resolution weed classification via convolutional neural network and superpixel based local binary pattern using remote sensing images. *Remote Sensing* **11**:1692.
- Feng X, Jiang Y, Yang X, Du M, Li X. 2019. Computer vision algorithms and hardware implementations: A survey. *Integration* **69**:309-320.
- Gebbers R, Adamchuk V I. 2010. Precision agriculture and food security. *Science* **327**: 828-831.
- Gerhards R, Andújar Sanchez D, Hamouz P, Peteinatos G G, Christensen S, Fernandez-Quintanilla C. 2022. Advances in site-specific weed management in agriculture—A review. *Weed Research* **62**:123-133.

- Gerhards R, Sökefeld M, Nabout A, Therburg R D, Kühbauch W. 2002. Online weed control using digital image analysis. *Z PflKrankh. PflSchutz, Sonderh.* **18**:421-427.
- Gharde Y, Singh P K, Dubey R P, Gupta P K. 2018. Assessment of yield and economic losses in agriculture due to weeds in India. *Crop Protection* **107**:12-18.
- Gianessi L P. 2013. The increasing importance of herbicides in worldwide crop production. *Pest management science* **69**:1099-1105.
- Goodfellow I, Bengio Y, Courville A. 2016. *Deep learning*. MIT press.
- Guerin L J, Guerin T F. 1994. Constraints to the adoption of innovations in agricultural research and environmental management: a review. *Australian journal of experimental agriculture* **34**:549-571.
- Hansen U S, Landau E, Patel M, Hayee B. 2021. Novel artificial intelligence-driven software significantly shortens the time required for annotation in computer vision projects. *Endoscopy International Open* **9**:621-626.
- He J, Li L, Xu J, Zheng C. 2018. ReLU deep neural networks and linear finite elements. arXiv preprint arXiv:1807.03973.
- Heckerman D, Wellman M P. 1995. Bayesian networks. *Communications of the ACM* **38**:27-31.
- Hsiao T Y, Chang Y C, Chou H H, Chiu C T. 2019. Filter-based deep-compression with global average pooling for convolutional networks. *Journal of Systems Architecture* **95**:9-18.
- Huang L, Joseph A D, Nelson B, Rubinstein B I, Tygar J D. 2011. Adversarial machine learning. Pages 43-58 in *Proceedings of the 4th ACM workshop on Security and artificial intelligence*.
- Ioffe S, Szegedy C. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Pages 448-456 in *International conference on machine learning*. PMLR.
- Kamasa J. 2021. Microchips: Small and Demanded. *CSS Analyses in Security Policy*, **295**.
- Kerkech M, Hafiane A, Canals R. 2018. Deep leaning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Computers and electronics in agriculture* **155**:237-243.
- Lahsen M. 2022. Wasting or Saving Lives: The Power of Chemicals and Laws. *Environment: Science and Policy for Sustainable Development* **64**:2-3.
- Lazonick W, Mazzucato M. 2013. The risk-reward nexus in the innovation-inequality relationship: who takes the risks? Who gets the rewards?. *Industrial and Corporate Change* **22**:1093-1128.
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* **521**:436-444.

- Li W, Sadigh D, Sastry S S, Seshia S A. 2014. Synthesis for human-in-the-loop control systems. Pages 470-484 in International Conference on Tools and Algorithms for the Construction and Analysis of Systems. Springer, Berlin, Heidelberg.
- Liakos K G, Busato P, Moshou D, Pearson S, Bochtis D. 2018. Machine learning in agriculture: A review. *Sensors* **18**:2674.
- Luo Y, Remillard J, Hoetzer D. 2010. Pedestrian detection in near-infrared night vision system. Pages 51-58 in 2010 IEEE Intelligent Vehicles Symposium. IEEE.
- Lutz W, Sanderson W, Scherbov S. 2001. The end of world population growth. *Nature* **412**:543-545.
- Madry A, Makelov A, Schmidt L, Tsipras D, Vladu A. 2017. Towards deep learning models resistant to adversarial attacks. arXiv preprint arXiv:1706.06083.
- McCarthy C L, Hancock N H, Raine S R. 2010. Applied machine vision of plants: a review with implications for field deployment in automated farming operations. *Intelligent Service Robotics* **3**:209-217.
- Miceli M, Schuessler M, Yang T. 2020. Between subjectivity and imposition: Power dynamics in data annotation for computer vision. *Proceedings of the ACM on Human-Computer Interaction* **4**:1-25.
- Miikkulainen R, Liang J, Meyerson E, Rawal A, Fink D, Francon O, ..., Hodjat B. 2019. Evolving deep neural networks. Pages 293-312 in Artificial intelligence in the age of neural networks and brain computing. Academic Press.
- Nawara J. 2010. Machine learning: face recognition technology evidence in criminal trials. *U. Louisville L. Rev.* **49**:601.
- Owen M D, Zelaya I A. 2005. Herbicide-resistant crops and weed resistance to herbicides. *Pest Management Science: formerly Pesticide Science* **61**:301-311.
- Pardey P G, Alston J M, Ruttan V W. 2010. The economics of innovation and technical change in agriculture. *Handbook of the Economics of Innovation* **2**:939-984.
- Pfeiffer D A. 2006. Eating fossil fuels: oil, food, and the coming crisis in agriculture. New Society Publishers.
- Potena C, Nardi D, Pretto A. 2016. Fast and accurate crop and weed identification with summarized train sets for precision agriculture. Pages 105-121 in International Conference on Intelligent Autonomous Systems. Springer, Cham.
- Pretty J. 2008. Agricultural sustainability: concepts, principles and evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363**:447-465.
- Prins A G, Eickhout B, Banse M, van Meijl H, Rienks W, Woltjer G. 2011. Global impacts of European agricultural and biofuel policies. *Ecology and society* **16**.

Recenti M, Ricciardi C, Gislason M, Edmunds K, Carraro U, Gargiulo P. 2019. Machine learning algorithms predict body mass index using nonlinear trimodal regression analysis from computed tomography scans. Pages 839-846 in Mediterranean Conference on Medical and Biological Engineering and Computing. Springer, Cham.

Redmon J, Farhadi A. 2018. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Redmon J, Divvala S, Girshick R, Farhadi A. 2016. You only look once: Unified, real-time object detection. Pages 779-788 in Proceedings of the IEEE conference on computer vision and pattern recognition.

Schaller N. 1993. The concept of agricultural sustainability. *Agriculture, ecosystems & environment* **46**:89-97.

Sharma R, Kamble S S, Gunasekaran A, Kumar V, Kumar A. 2020. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research* **119**:104926.

Spelda P, Stritecky V. 2021. Human Induction in Machine Learning: A Survey of the Nexus. *ACM Computing Surveys (CSUR)* **54**:1-18.

Timmermann C, Gerhards R, Kühbauch W. 2003. The economic impact of site-specific weed control. *Precision Agriculture* **4**:249-260.

Wang K, Gao X, Zhao Y, Li X, Dou D, Xu C Z. 2019. Pay attention to features, transfer learn faster CNNs. in International conference on learning representations.

Zhang W, Hansen M F, Volonakis T N, Smith M, Smith L, Wilson J, Ralston G, Broadbent L, Wright G. 2018. Broad-leaf weed detection in pasture. Pages 101-105 in 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). IEEE.