



Fakulta zemědělská
a technologická
Faculty of Agriculture
and Technology

Jihočeská univerzita
v Českých Budějovicích
University of South Bohemia
in České Budějovice

JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH FAKULTA ZEMĚDĚLSKÁ A TECHNOLOGICKÁ

Katedra techniky a kybernetiky

Bakalářská práce

Příprava datasetů termosnímků paznehtů skotu vhodných pro
trénování CNN

Autor práce: Radek Zahajský

Vedoucí práce: Ing. Mgr. Radim Stehlík

České Budějovice
2023

Prohlášení

Prohlašuji, že jsem autorem této kvalifikační práce a že jsem ji vypracoval pouze s použitím pramenů a literatury uvedených v seznamu použitých zdrojů.

V Českých Budějovicích dne

.....

Podpis

Abstrakt

Bakalářská práce zabývající se přípravou datasetu pro trénování CNN, počítačovým viděním a hlubokým učením neuronové sítě. Dále důležitý aspekt tvoří využití této technologie, jednotlivé aplikace, formáty a barevné systémy. Praktická část popisuje postup při anotování snímků v aplikaci CVAT. Při práci se postupuje získáváním jednotlivých informací o tématech a porovnáváním aplikací a formátů. Hlavním výsledkem práce je vytvoření kvalitního datasetu pro identifikaci paznehtů dojnic.

Klíčová slova: anotace, dataset, počítačové vidění, pazneht, skot

Abstract

Bachelors work dealing with dataset preparation for CNN training, computer vision and deep neural network learning. Other important aspects are the use of this technology, the different applications, formats and color spaces. The practical part describes the process of image annotation in CVAT software. The work involves individual extraction of information on each subject and comparison of applications and formats. The main result of the work is the creation of a high quality dataset for the identification of dairy cow hooves.

Keywords: annotation, dataset, computer vision, hoof, cattle

Poděkování

Chtěl bych poděkovat vedoucímu práce Ing. Mgr. Radimu Stehlíkovi, za nasměrování, jak vést bakalářskou práci, a dále za jeho vstřícné vedení a spolupráci, trpělivost a cenné informace. Dále bych chtěl poděkovat své rodině za neúnavnou podporu a trpělivost při zhotovování bakalářské práce.

Obsah

Úvod.....	7
1 Teoretická část	8
1.1 Počítačové vidění	8
1.2 Konvoluční neuronová síť	11
1.3 Učení neuronové sítě	13
1.4 Využití technologií v zemědělství.....	14
1.5 Software pro anotaci.....	16
1.5.1 CVAT.....	16
1.5.2 V7.....	18
1.5.3 VGG Image Annotator.....	19
1.5.4 LabelMe	21
1.5.5 LabelImg	22
1.5.6 Label Studio	23
1.6 Anotační formáty a datasety.....	25
1.6.1 MS COCO.....	25
1.6.2 PASCAL VOC	28
1.6.3 YOLO.....	30
1.7 Veřejně dostupné datasety.....	31
1.8 Barevné systémy.....	33
1.8.1 RGB	34
1.8.2 CMYK.....	35
1.8.3 HSV.....	35
1.8.4 L*a*b*	36
1.8.5 Termální snímky	38
2 Cíl práce	41
3 Metodika	42

3.1	Vstupní data.....	42
3.2	Příprava dat.....	43
3.3	Pracovní prostředí.....	45
3.4	Anotace dat.....	49
3.5	Export dat	51
4	Výsledky a diskuse.....	53
	Závěr	57
	Seznam použité literatury.....	58
	Seznam obrázků	64
	Seznam tabulek	65
	Seznam grafů.....	66
	Seznam použitých zkratk.....	67

Úvod

Bakalářská práce se zabývá přípravou datasetu termálních snímků paznehtů skotu pro trénování konvoluční neuronové sítě, jenž v dnešním světě počítačového vidění umožňuje strojům rozpoznávat objekty na fotografiích, které slouží jako vstupní informace pro učení neuronové sítě. Konvoluční neuronové sítě zpracovávají nejrůznější vstupní data a díky této technologii je přístroj schopen rozeznat objekty na základě široké databáze, která obsahuje stejné objekty z jiného úhlu pohledu.

Dopravní, zdravotnický, a především zemědělský průmysl této technologie využívá pro zdokonalení postupů a získání lepších výsledků. U skotu je důležité udržovat celkové zdraví pro ovlivnění požadované užitkovosti. Zejména nemocné paznehty mohou zapříčinit nízký pohyb skotu a tím i snížení produkčních schopností. Infekce, záněty a jiná onemocnění je nutno určitým způsobem hlídat, a právě technologie počítačového vidění může za pomoci termálních snímků tyto problémy odhalit. Pro řešení těchto problémů musí neuronová síť obsahovat velkou databázi objektů, díky kterým je schopna se naučit rozeznávat požadované předměty zkoumání.

V práci jsou srovnány vybrané aplikace pro anotaci objektů na snímku, z nichž každá má své výhody a nevýhody, které hrají klíčovou roli ve výběru pro práci uživatele se snímky. Formáty, ve kterých jsou anotace vytvářeny mají široké zastoupení, z nichž každý pracuje na jiné bázi.

Kamery a termokamery využívají barevné systémy pro zachycení a následné vyobrazení snímků. Barevných palet existuje nespočet, nicméně tato práce se zabývá podstatnějšími z nich. Příprava datasetu je klíčovým faktorem v oboru počítačového učení. Preciznost při práci se snímky a kvalita anotací je nesmírně důležitá pro získání požadovaných výsledků.

1 Teoretická část

Počítačové vidění a hluboké učení jsou dvě nejzajímavější a nejrychleji rozvíjející se oblasti umělé inteligence. Již nyní mají významný dopad na celou řadu průmyslových odvětví, včetně zdravotnictví, automobilového průmyslu a volnočasových aktivit, a mají potenciál v budoucnu způsobit revoluci v mnoha dalších oborech (Šonka et al., 2013; Goodfellow et al., 2016).

1.1 Počítačové vidění

Počítačové vidění je obor umělé inteligence a počítačové vědy, který se zaměřuje na to, aby stroje mohly interpretovat a chápat vizuální data z okolního světa. Může se jednat o obrázky, videa či 3D skeny a zahrnuje širokou škálu využití, včetně rozpoznávání objektů, porozumění scéně, analýzy obrazu, videa a rozšířené a virtuální reality. Jednou z nejvýznamnějších hnacích sil růstu počítačového vidění je rostoucí dostupnost digitálních dat, včetně obrázků a videí, a také vývoj pokročilých výpočetních technologií, jako je hluboké učení, které dokáže rychle a přesně zpracovávat a analyzovat velké množství dat (Šonka et al., 2013).

Počítačové vidění lze aplikovat v celé řadě průmyslových odvětví, včetně zdravotnictví, bezpečnosti, dopravy a zemědělství. Co se týče zdravotnictví, lze technologie počítačového vidění využít k analýze lékařských snímků, odhalování abnormalit a pomoci při diagnostice. V dopravě lze počítačové vidění využít pro autonomní vozidla k detekci a reakci na překážky a dopravní signály. V oblasti volnočasových aktivit lze počítačové vidění využít k vytváření pohlcujících zážitků v rozšířené a virtuální realitě (Kolář, 2020).

Detekce objektů je jednou z nejběžnějších úloh počítačového vidění, která je založena na principu rozeznávání a identifikaci objektů na obrázcích a videích. To lze využít pro účely takových úloh, jako je detekce a sledování osob nebo vozidel na záznamech z bezpečnostních kamer, identifikace předmětů v maloobchodním prostředí nebo rozpoznávání obličejů na platformách sociálních médií (Forsyth a Ponce, 2011).

Porozumění konkrétní podobě obrazu je další oblastí počítačového vidění, která se zaměřuje na zpracování obsahu a kontextu vizuálních objektů. Jedná se o analýzu snímků nebo videí za účelem pochopení vztahů mezi objekty a jejich okolím, kterou lze využít v aplikacích, jako je autonomní navigace, monitorování životního prostředí a plánování měst (Russ, 2002).

Analýza je širší obor, který zahrnuje řadu technik a metod pro zpracování a získávání informací z vizuálních dat. Může zahrnovat techniky, jako je segmentace obrazu, získávání vlastností a sledování pohybu, a může být použita pro aplikace, jako je komprese obrazu a videa nebo vyhledávání obrazu na základě obsahu. Jednou z klíčových výzev v oblasti počítačového vidění je vývoj algoritmů a modelů, které dokáží přesně interpretovat a pochopit vizuální data, která mohou být složitá, zašuměná a proměnlivá. Hluboké učení, podobor strojového učení, se díky své schopnosti učit složitá data z velkých souborů dat stalo mocným nástrojem pro vývoj pokročilých modelů počítačového vidění (Forsyth a Ponce, 2011).

Navzdory významnému pokroku, kterého bylo v posledních letech dosaženo v oblasti počítačového vidění a hlubokého učení, stále existuje mnoho výzev a omezení, které je třeba řešit. Patří mezi ně otázky, jako je potřeba velkého množství vysoce kvalitních trénovacích dat, riziko zkreslení a nesprávnosti modelů a potřeba, aby modely byly odolné vůči změnám osvětlení, pozorovacím úhlům a dalším faktorům prostředí. Potenciální přínosy počítačového vidění jsou však značné a s dalším rozvojem a zdokonalováním této technologie mohou způsobit revoluci v celé řadě průmyslových oborů. Počítačové vidění mění způsob, jakým komunikujeme s okolním světem. Otevírá nové možnosti pro inovace a objevy od zdravotnictví přes dopravu až po volnočasové aktivity (Šonka et al., 2013).

Jak již bylo zmíněno, tento obor využívá obraz a video k detekci, klasifikaci a sledování objektů nebo událostí za účelem pochopení reálné situace. Úlohy počítačového vidění lze řešit různými přístupy, od porovnávání vzoru se snímkem až po metodu deep learning, nabízející velkou přesnost a spolehlivost. Techniky zpracování obrazu lze rozdělit do dvou kategorií:

- Klasické zpracování obrazu – Zaměřuje se na tradiční úlohy, jako je denoizace a filtrace obrazu, úprava kontrastu, manipulace s barevným systémem nebo měření a transformace obrazu. Kromě zpracování a předzpracování obrazu zahrnují klasické techniky zpracování obrazu detekci objektů na základě barevné informace, morfologické operace, segmentaci a vytváření masek.
- Počítačové vidění – přináší do rozboru obrazu hlubší rozměr se zaměřením na detekci, klasifikaci a sledování objektů na základě jejich charakteristických rysů (Šonka et al., 2013).

První skupina úloh počítačového vidění spočívá ve vyhledávání vzorových objektů v obrázku nebo videosekvenci. To se provádí pomocí algoritmů, které hledají podobné rysy v obraze a hledaném vzoru. Výsledkem může být poloha hledaného objektu ve scéně nebo transformace, která mapuje vzor na snímek scény (Forsyth a Ponce, 2011).

Porovnávají se podobnosti zjištěné na snímku a na modelu. Poloha shodných dvojic podobností určuje polohu vzorového objektu v obraze. Vyhledávání i porovnávání jsou plně automatizované a pro tuto funkci existuje mnoho uživatelsky definovaných parametrů (Kolář, 2020).

Kromě již popsaného vyhledávání modelových objektů je možné detekovat i obecnější kategorie objektů, včetně některých odrůd. Tyto detektory musí být nejprve vycvičeny na konkrétním typu objektu pomocí sady vzorových dat a poté je lze použít na nové snímky (Russ, 2002).

Sledování objektů není jen opakovaná detekce objektů aplikovaná na jednotlivé snímky videa. Při sledování se odborníci snaží najít souvislosti a změny scény v sobě jdoucích snímcích (Forsyth a Ponce, 2011).

Klasifikace snímaných objektů je jiný typ úlohy než vyhledávání konkrétních objektů v obraze. Při klasifikaci je celému snímku přiřazena jedna klasifikační kategorie, která charakterizuje snímaný objekt. Úloha se může týkat objektů typických pro okolní svět nebo může být specializovaná na rozpoznávání vybraných objektů z dané skupiny (Forsyth a Ponce, 2011).

Možnosti využití strojového vidění:

Počítání lahví v přepravce či kontrola tablet v blistru, kontrola správnosti nasazení zátky, kontrola úplnosti a správnosti obsahu před uzavřením. Určení správné aplikace barvy či nátěru na výrobku, třídění typů dílů dle barvy, ověření správného materiálu výrobku. Identifikace a verifikace čárového nebo datamatrixového kódu, identifikace dílů dle nápisu, zajištění nezaměnitelnosti obalu k obsahu. Rozpoznání správného dílu pomocí nápisu, kódu či charakteristického znaku, kontrola správné polohy či natočení dílu. Kontrola krimpování vodičů, kontrola správné montáže sad, kontrola orientace štítku. Přesné měření a kontrola tolerancí se zaměřuje na velikost dílu, průměr otvoru, rozteč otvorů, závity, úhel zúžení, vzdálenost hran a tak dále. Poškození hran, jejich promáčknutí, oděry od barvy a jiné povrchové závady, poškození pečeti nebo písmen, nesprávný potisk, kontrola bublin ve skle nebo plastu a kontrola stavu

naneseného lepidla nebo tmelu. Oděrky po řezání nebo lisování, nečistoty po obrábění dílů, nedokončené nebo přetmelené díly vstřikovaných dílů. (Šonka et al., 2013).

1.2 Konvoluční neuronová síť

Konvoluční neuronové sítě (CNN) jsou třídou umělých neuronových sítí, které se hojně využívají v aplikacích počítačového vidění. Jsou navrženy tak, aby zpracovávaly vstupní data, která mají strukturu podobnou mřížce, jako jsou obrázky a videa, extrahováním relevantních rysů z dat prostřednictvím procesu zvaného konvoluce (Aghdam a Heravi, 2017).

Sítě byly poprvé představeny v 90. letech 20. století, ale teprve na počátku roku 2010 se začaly široce používat v úlohách počítačového vidění. Jedním z klíčových faktorů, které vedly k rozmachu CNN, byla dostupnost velkých datových sad, jako je ImageNet, která obsahovala miliony označených obrázků. To umožnilo výzkumníkům trénovat rozsáhlé a složité neuronové sítě, které dokázaly s vysokou přesností rozpoznávat vzory a rysy v obrazech (Hassaballah a Awad, 2020).

Základní složkou CNN je konvoluční vrstva. Ta se skládá ze sady filtrů, z nichž každý je zodpovědný za detekci určitého rysu ve vstupních datech. Jeden filtr může detekovat hrany, zatímco jiný rohy. Proces konvoluce spočívá v posouvání filtru po vstupních datech, přičemž se vypočítá bodový součin mezi filtrem a částí vstupu, na které je právě překryt. Výsledkem je mapa prvků, která zvýrazňuje ty části vstupních dat, jež jsou pro prvky detekované filtrem nejrelevantnější. Výstupem konvoluční vrstvy je sada map prvků, z nichž každá odpovídá jinému filtru (Hassaballah a Awad, 2020).

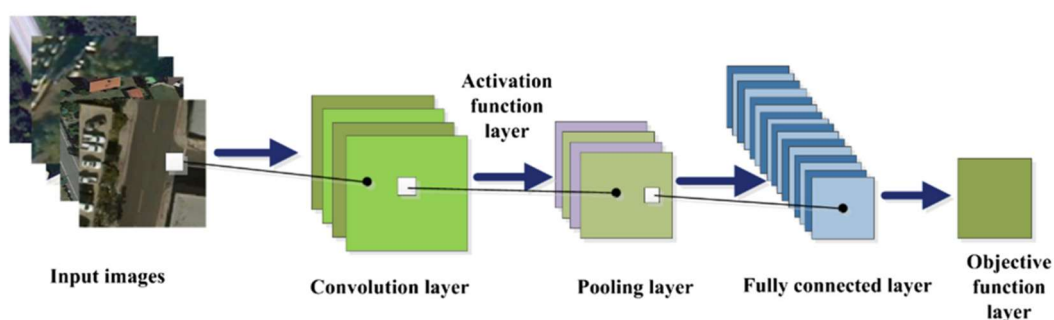
Kromě konvolučních vrstev CNN často obsahují i vrstvy sdružující, které snižují velikost map prvků a zároveň zachovávají nejdůležitější informace. Nejběžnějším typem sdružování je takzvaný max pooling, který vybírá maximální hodnotu z každé malé oblasti mapy prvků, čímž efektivně snižuje velikost vzorku mapy prvků. Jakmile jsou mapy prvků vytvořeny konvolučními a poolovacími vrstvami, jsou předány jedné nebo více plně propojeným vrstvám, které provádějí klasifikaci nebo regresi na základě prvků extrahovaných ze vstupních dat. Výstupem poslední vrstvy je předpovězená třída nebo hodnota (Aghdam a Heravi, 2017).

Jednou z klíčových výhod CNN je jejich schopnost automaticky se učit relevantní rysy ze vstupních dat, aniž by bylo nutné ruční vytváření rysů. Hluboké učení je podoblast strojového učení, která se v posledních letech stává stále populárnější díky své

schopnosti učit se komplexní reprezentaci dat z velkých souborů datasetů. Hluboké učení mělo významný dopad zejména na oblast počítačového vidění, kde bylo použito k vývoji výkonných modelů pro analýzu a klasifikaci obrazu. Díky tomu jsou CNN vysoce flexibilní a přizpůsobitelné různým úlohám a datovým sadám. Kromě toho si poradí i se vstupními daty, která jsou velmi proměnlivá, jako jsou obrázky s různými světelnými podmínkami, úhly pohledu a pozadím (Goodfellow et al., 2016; Aghdam a Heravi, 2017).

CNN se používají v široké škále aplikací, včetně klasifikace obrazu, detekce objektů a rozpoznávání obličejů. Dosáhly nejlepších výsledků v mnoha srovnávacích testech počítačového vidění a stále častěji se používají v průmyslových odvětvích, jako jsou samořízená auta a analýza lékařských snímků. Mají však i některá omezení. K trénování vyžadují velké množství označených dat a jejich učení a provoz mohou být výpočetně nákladné, zejména v případě velkých souborů dat. Navíc nejsou vždy transparentní ve svém rozhodování, což může být problém v oborech, kde je důležitá vysvětlitelnost, například ve zdravotnictví (Hassaballah a Awad, 2020).

Přestože se v tomto oboru vyskytují jistá omezení, CNN sítě změnily obor počítačového vidění a nadále jsou hlavním prvkem pokroku v analýze obrazu a videa. Díky pokračujícímu výzkumu nových struktur a optimalizačních metod je pravděpodobné, že budou i nadále hrát klíčovou roli v budoucnosti umělé inteligence a počítačového vidění. Na Obrázku 1.1 (Liu et al., 2019) je vyobrazena klasická struktura CNN (Aghdam a Heravi, 2017; Hassaballah a Awad, 2020).



Obrázek 1.1: Struktura CNN

1.3 Učení neuronové sítě

Jednou z klíčových výzev při vývoji modelů hlubokého učení pro analýzu obrazu je nutné velké množství anotovaných dat potřebných pro trénování. Anotovanými daty se rozumí obrazy, které byly ručně označeny informacemi o objektech nebo vlastnostech přítomných v obraze, jako jsou hranice objektů, třídy objektů nebo segmentační masky (Liu et al., 2019).

Proces anotace může být nejen časově, ale také pracovní náročný, jelikož vyžaduje, aby člověk pečlivě prozkoumal každý obrázek a označil jej příslušnými informacemi. Kvalita a množství anotovaných dat jsou však klíčové pro trénování přesných a efektivních modelů hlubokého učení. Pro obrazová data lze použít několik různých typů anotací, z nichž každá má své výhody a nevýhody. Jedním z běžných přístupů je použití ohraničujících boxů, konkrétně obdélníkových, které obklopují objekt zájmu na obrázku. Anotace ohraničujících polí jsou relativně jednoduché a snadno se vytvářejí, ale bohužel poskytují jen omezené informace o tvaru a orientaci objektu (Goodfellow et al., 2016).

Další typ anotací, který spočívá v označení každého pixelu v obrázku značkou kategorie nebo třídy je důležitý pro semantickou segmentaci. Tyto anotace poskytují podrobnější informace o tvaru a hranicích objektu, ale jejich generování může být obtížnější a časově náročnější než generování ohraničujících rámečků (Goodfellow et al., 2016).

Anotace pomocí polygonálních masek, které bývají aplikovány pro segmentaci instancí, spočívají v identifikaci a označení jednotlivých objektů v obraze, nikoli pouze v klasifikaci pixelů. Tyto segmentace jsou užitečné pro aplikace, jako je detekce a sledování objektů, ale jejich generování je ještě obtížnější než sémantické segmentace. Modely hlubokého učení lze trénovat pomocí různých typů anotací v závislosti na konkrétní úloze a aplikaci. Například model pro detekci objektů může být trénován pomocí ohraničujících boxů, zatímco model pro sémantickou segmentaci může být trénován pomocí anotací na úrovni pixelů (Goodfellow et al., 2016).

Pro rozšíření dat lze aplikovat metody augmentace sloužící ke generování dalších trénovacích dat použitím transformací původních obrázků, jako je rotace, škálování a převrácení (Hassaballah a Awad, 2020).

Pro rychlejší získání optimálních výsledků je možné použít přenosové učení, což je technika, která zahrnuje použití předem natrénovaných modelů hlubokého učení jako výchozího bodu pro novou úlohu namísto trénování modelu od začátku. To může

být užitečné, pokud jsou k dispozici jen omezená anotovaná data nebo pokud je cílová úloha podobná úloze, na které byl původně trénován předtrénovaný model (Hassaballah a Awad, 2020).

Předtrénované modely se obvykle učí na rozsáhlých sadách obrazových dat, jako je například ImageNet, která obsahuje miliony označených obrázků napříč tisíci kategoriemi. Smyslem předtrénovaných modelů je využít vlastnosti naučené modelem trénovaným na jedné úloze a přenést je na novou úlohu s podobnými vlastnostmi. Například předtrénovaný model vycvičený na úlohu klasifikace obrázků lze použít jako výchozí bod pro novou úlohu, jako je detekce objektů nebo sémantická segmentace (Hassaballah a Awad, 2020).

Modely se obvykle trénují pomocí již zmíněných hlubokých konvolučních neuronových sítí (CNN), které jsou určeny k učení hierarchických vlastností ze surových pixelových dat. Tyto sítě mohou být velmi rozsáhlé, s desítkami nebo dokonce stovkami vrstev, což způsobuje, že jejich trénování od nuly je výpočetně velmi nákladné (Hassaballah a Awad, 2020).

Použitím předtrénovaného modelu jako výchozího bodu mohou výzkumní pracovníci a odborníci z praxe ušetřit čas a výpočetní zdroje, protože takovýto model již obsahuje naučené funkce, které lze použít jako výchozí bod pro novou úlohu. Tento postup se nazývá transfer learning a je oblíbenou technikou v oblasti deep learningu, která se zabývá omezenými daty a výpočetními zdroji (Goodfellow et al., 2016).

Pro použití v hlubokém učení je k dispozici několik populárních modelů, včetně modelů VGGNet, ResNet, Inception a Xception. Tyto modely byly předem natrénovány na rozsáhlých sadách obrazových dat a dosáhly špičkového výkonu v různých úlohách analýzy obrazu. Kromě předtrénovaných modelů pro analýzu obrazu jsou k dispozici také modely pro jiné typy dat, například pro zpracování přirozeného jazyka (NLP) a rozpoznávání řeči (Rosebrock, 2017).

1.4 Využití technologií v zemědělství

Technologie působící v oborech počítačového vidění a hlubokého učení mají potenciál způsobit revoluci v živočišné výrobě a zemědělství tím, že zlepší efektivitu a produktivitu podniku a životní podmínky zvířat. Tyto technologie mohou poskytnout vhled do chování, zdraví a výživy zvířat a umožnit přesnější a cílenější zásahy (Fák, 2014; Vseoprmyslu.cz, 2020).

V živočišné výrobě lze tyto technologie využít ke sledování a analýze chování zvířat, což může poskytnout poznatky o jejich zdraví a dobrých životních podmínkách. Algoritmy počítačového vidění mohou například detekovat změny v postoji nebo chůzi zvířete, jenž mohou indikovat kulhání či jiné zdravotní problémy. Taktéž mohou sledovat chování při krmení a identifikovat změny v příjmu krmiva, což by mohlo být včasným příznakem nemoci nebo nedostatku výživy (Mangla, 2022; Patel, 2022).

Počítačové vidění a hluboké učení lze využít také v precizním chovu hospodářských zvířat, kde se senzory a kamery používají ke sledování chování zvířat, podmínek prostředí a parametrů produkce. To může zemědělcům pomoci optimalizovat spotřebu krmiva a vody, odhalit ohniska nemocí a předcházet jim a zlepšit celkovou produktivitu a efektivitu provozu (Patel, 2022).

Dále lze také sledovat růst a zdraví plodin, identifikaci napadení škůdci a optimalizaci zavlažování a hnojení. Algoritmy počítačového vidění mohou analyzovat satelitní snímky nebo snímky z dronů a odhalit změny ve vegetaci, které by mohly indikovat stres nebo onemocnění. Tyto informace lze využít k cílení zásahů, jako je aplikace pesticidů nebo úprava zavlažovacích plánů, na konkrétní oblasti pole namísto rovnoměrného ošetření celé plochy pole. Kromě toho lze obě technologie využít k automatizaci úkolů, které se v současné době provádějí ručně, jako je sklizeň ovoce a zeleniny, což může zvýšit efektivitu a snížit náklady na pracovní sílu a zároveň snížit riziko poškození plodin a tím zlepšit jejich kvalitu (Fák, 2014).

Jednou z hlavních výhod neuronových sítí v živočišné výrobě a zemědělství celkově je jejich schopnost rychle a přesně zpracovávat velké množství dat, jenž umožňuje sledování a rozhodování v reálném čase, což může být v rychlém a dynamickém prostředí klíčové (Mangla, 2022).

Na druhou stranu zavádění těchto technologií do živočišné výroby a zemědělství celkově přináší řadu výzev. Za hlavní výzvu lze považovat potřebu vysoce kvalitních dat, které může být v některých podmínkách obtížné získat. Kupříkladu venkovní prostředí je pro kamery a senzory z hlediska dohledu a počasí náročné, čímž se zhoršuje i možnost zvířata adekvátně sledovat a monitorovat (Mangla, 2022).

Jako další výzvu je lze považovat potřebu specializovaných odborných znalostí a infrastruktury, jejichž vývoj může být nákladný a časově náročný. To zahrnuje potřebu vysoce výkonných výpočetních zdrojů a specializovaného softwaru a algoritmů (Vseoprumsly.cz, 2020).

Tyto výzvy mohou být náročné, avšak potenciální přínosy jsou v živočišné výrobě a zemědělství značné. Vzhledem k tomu, že se tyto technologie dále vyvíjejí a stávají se dostupnějšími, mají potenciál změnit toto odvětví, zvýšit efektivitu, produktivitu a udržitelnost a zároveň zlepšit životní podmínky zvířat a snížit dopad na životní prostředí (Vseoprumsu.cz, 2020).

1.5 Software pro anotaci

Anotační software je typ programu, který se používá k označování nebo anotování dat způsobem srozumitelným pro algoritmy strojového učení. Běžně se používá v oblasti počítačového vidění, kde se anotují rozsáhlé soubory dat obrázků nebo videí, aby bylo možné trénovat modely strojového učení pro rozpoznávání a klasifikaci objektů, detekci vzorů nebo provádění jiných úkolů. Obecně anotační software umožňuje uživatelům ručně označit nebo anotovat konkrétní prvky v datech, jako jsou objekty, oblasti nebo body zájmu. Tyto informace jsou pak použity k trénování modelů strojového učení, což jim umožňuje tyto prvky automaticky identifikovat nebo rozpoznat. Software může také obsahovat nástroje pro automatické generování anotací na základě předem definovaných pravidel (Boesch, 2023).

Existuje mnoho různých typů anotačních softwarů, od jednoduchých webových nástrojů až po složité specializované softwarové balíčky. Některé mohou být speciálně navrženy pro určité typy dat, jako jsou lékařské snímky RTG a USG nebo satelitní snímky Země. Jiný software může nabízet řadu nástrojů, jako jsou ohraničující boxy, segmentace a bodové anotace, které podporují širokou škálu případů použití (Boesch, 2023).

Takovýto software hraje zásadní roli při dosažení kvalitních výsledků modelů strojového učení v široké škále aplikací a dovoluje uživatelům označovat a anotovat rozsáhlé datové soubory obrázků nebo videí, čímž poskytuje trénovací data potřebná k vývoji přesných a efektivních modelů strojového učení (Boesch, 2023).

1.5.1 CVAT

CVAT (Computer Vision Annotation Tool) je softwarový nástroj s otevřeným zdrojovým kódem vyvinutý společností Intel, jenž usnadňuje anotování obrázků a videí pro účely trénování modelů počítačového vidění. Nástroj je navržen tak, aby zjednodušil proces anotace tím, že poskytuje intuitivní rozhraní, které uživatelům umožňuje rychle a přesně anotovat obrázky a videa (Syndorenko, 2021).

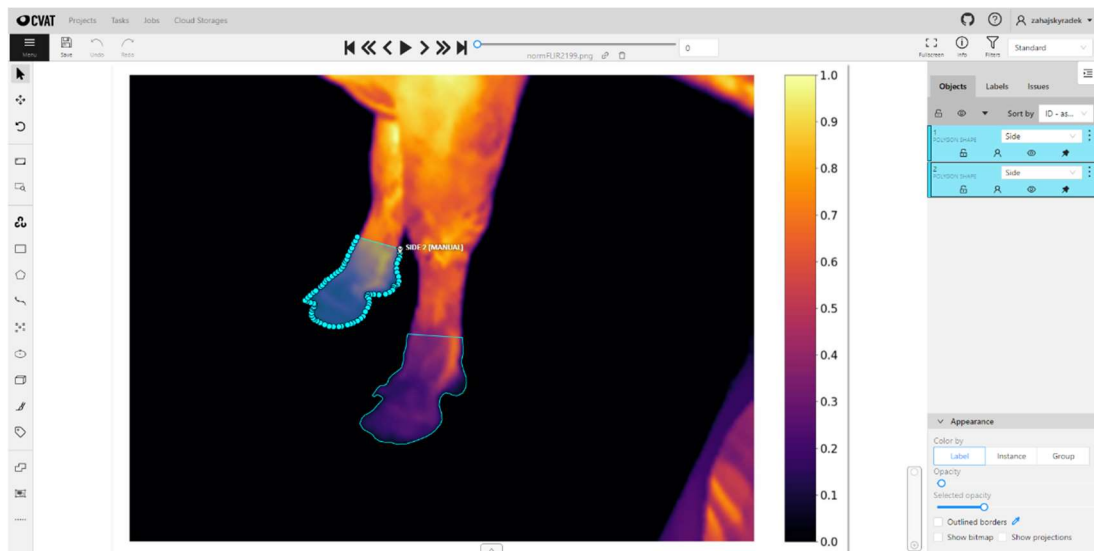
Software poskytuje řadu anotačních prvků, včetně ohraničujících boxů, polygonů, čar a bodů, které uživatelům umožňují přesně anotovat objekty na obrázcích a videích. Software také poskytuje možnosti spolupráce, které umožňují více uživatelům pracovat na stejném anotačním projektu současně. To poskytuje distribuovaným týmům zpracovávat anotační úlohy společně a napomáhá zefektivnit anotační proces tím, že poskytuje více lidem pracovat na různých částech téhož projektu (Boesch, 2023).

Další klíčovou vlastností systému CVAT je jeho přizpůsobitelné rozhraní. Nástroj dovoluje uživatelům přizpůsobit si rozhraní podle svých preferencí a pracovních postupů, což z něj činí vysoce flexibilní nástroj. Toto přizpůsobení může pomoci zvýšit produktivitu a přesnost, protože uživatelé mohou nástroj přizpůsobit svým specifickým potřebám. Obsahuje také funkce pro správu projektů, které uživatelům umožňují spravovat více anotačních projektů současně. To poskytuje centralizované místo pro ukládání a správu anotačních dat a pomáhá zajistit, aby anotační data byla uspořádaná a snadno dostupná (Boesch, 2023).

CVAT podporuje řadu exportních formátů, včetně COCO, Pascal VOC a YOLO. Tato nabídka uživatelům umožňuje snadno exportovat anotovaná data pro použití při trénování modelů počítačového vidění. To napomáhá integraci anotovaných dat do stávajících řetězců strojového učení (Solawetz, 2022).

Ačkoli má mnoho výhod, je potřeba vzít v úvahu i některé nevýhody. Jednou z potenciálních nevýhod je učení spojené s tímto nástrojem. Prostředí CVATu může být složité na naučení, zejména pro uživatele, kteří s počítačovým viděním a anotováním obrazu začínají. Dále poskytuje řadu anotačních nástrojů, nicméně nemá pokročilé funkce detekce objektů, což může omezit jeho užitečnost pro některé aplikace (Solawetz, 2022).

Celkově lze CVAT považovat za výkonný nástroj pro anotaci obrázků a videí pro obor počítačového vidění. Díky řadě funkcí a možnosti přizpůsobení je flexibilním nástrojem pro širokou škálu anotačních úloh. Jeho složitost a náročnost na zdroje jej však mohou činit méně vhodným pro uživatele s omezenými zkušenostmi v oblasti počítačového vidění a anotování obrázků (Syndorenko, 2021). Prostředí softwaru je vyobrazeno na Obrázku 1.2 (autor).



Obrázek 1.2: Pracovní prostředí v softwaru CVAT

1.5.2 V7

V7 je webový anotační softwarový nástroj, který je navržen tak, aby zjednodušil a zefektivnil proces anotování velkých souborů dat obrázků a videí (Humansintheloop.org, 2020).

Hlavní vlastností V7 je jeho flexibilita. Nástroj podporuje širokou škálu anotačních úloh, včetně ohraničujících boxů, polygonů, čar a bodů, takže je vhodný pro širokou škálu případů použití. Poskytuje navíc řadu možností přizpůsobení, včetně možnosti vytvářet vlastní anotační třídy a přizpůsobovat vzhled rozhraní podle preferencí uživatele (Humansintheloop.org, 2020).

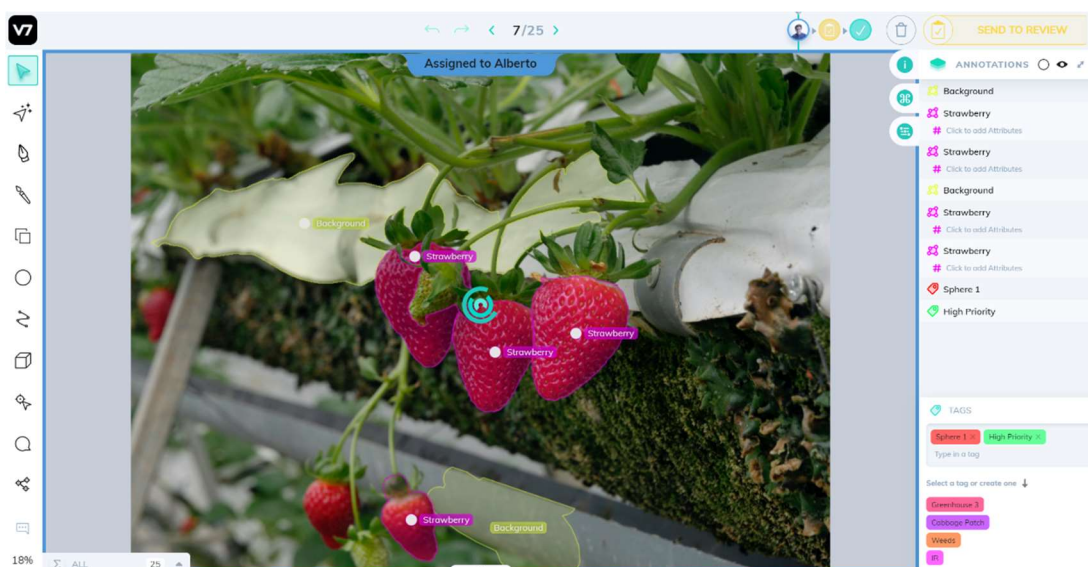
Další významnou vlastností systému jsou možnosti spolupráce. Nástroj podporuje práci více uživatelů na stejném anotačním projektu současně, což může přispět ke zvýšení efektivity a přesnosti. Kromě toho poskytuje řadu nástrojů pro správu anotačních projektů, včetně možnosti sledovat průběh, spravovat uživatele a vytvářet zprávy. V7 obsahuje také pokročilé funkce pro podporu kontroly kvality a správy dat. Nástroj poskytuje nástroje pro ověřování a opravu anotací, čímž zajišťuje, že anotace jsou přesné a konzistentní. Kromě toho obsahuje funkce pro správu velkých souborů dat obrázků a videí, včetně podpory verzování, rozšiřování a vzorkování dat. Jednou z výhod tohoto systému je jeho snadné používání (Lunden, 2022).

Software má intuitivní, uživatelsky přívětivé rozhraní viz Obrázek 1.3 (v7labs.com, 2023), které uživatelům usnadňuje začít s anotačními úlohami. V7 navíc poskytuje řadu školení a podpůrných zdrojů, které uživatelům pomohou nástroj co nejlépe využít. Podporuje také většinu široce využívaných typů souborů, vlastnosti

snímků a formátů. Mezi ně patří například ultra vysoké rozlišení, multispektrální formáty a formát PDF (Lunden, 2022).

Při používání softwaru je však třeba uvážit i některé potenciální nevýhody. Nástroj je webový, což znamená, že k jeho používání je nutné připojení k internetu (Yao, 2020).

V7 je výkonný a flexibilní anotační softwarový nástroj, který lze použít k anotování velkých datasetů obrázků a videí pro použití jako zdroje strojového učení. Jeho snadné použití, možnosti spolupráce a pokročilé funkce z něj činí cenný nástroj pro anotaci a správu dat. Uživatelé by však měli pečlivě vyhodnotit své specifické potřeby a požadavky, než se rozhodnou používat V7, jelikož je volně dostupný pouze pro edukační potřeby. Pro organizace do deseti členů je zdarma zkušební verze, která je bohužel při delším vyžívání zpoplatněná. Pro větší skupiny uživatelů je potřeba sjednat koupi (Yao, 2020).



Obrázek 1.3: Pracovní prostředí softwaru V7

1.5.3 VGG Image Annotator

VGG Image Annotator (VIA) je populární open-source anotační software, který se hojně používá v oblasti počítačového vidění. Je výhradně založen na kódech HTML, Javascriptu a CSS (Dutta, a Zisserman, 2019).

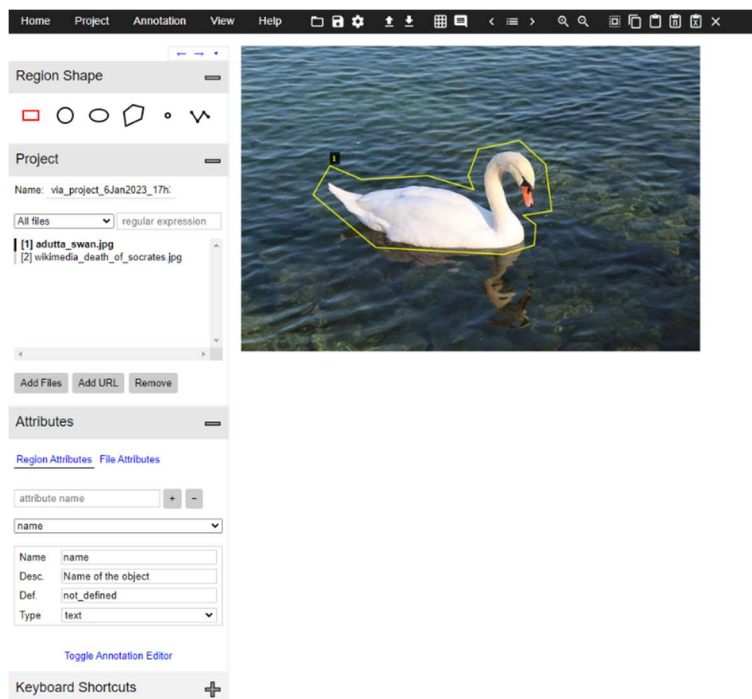
VIA je navržen tak, aby zjednodušil proces anotování velkých souborů dat a umožnil uživatelům rychle a přesně označovat obrázky pro použití v modelech strojového učení a dalších oborech (Aljabri et al., 2022).

Podporuje širokou škálu anotačních úloh a disponuje stejnými nástroji jako například software V7. Nabízí velkou řadu možností přizpůsobení, včetně možnosti vytvářet vlastní anotační třídy a upravovat vzhled anotačního rozhraní podle preferencí uživatele. Velkou výhodou je snadné používání. Nástroj má jednoduché, uživatelsky přívětivé rozhraní, jenž uživatelům usnadňuje začít s anotačními úlohami. VIA navíc poskytuje školení a podpůrné zdroje, které uživatelům pomohou software co nejlépe využít (Dutta, a Zisserman, 2019; Aljabri et al., 2022).

VIA obsahuje pokročilé funkce pro podporu kontroly kvality a správy dat a nástroje pro ověřování a opravu anotací, díky kterým zajišťuje, že anotace jsou přesné a konzistentní. Kromě toho nabízí funkce pro správu velkých souborů dat snímků, včetně podpory určování verzí, rozšiřování a vzorkování dat. Významnou výhodou je otevřený zdrojový kód. Nástroj je volně ke stažení a použití a jeho zdrojový kód je k dispozici na serveru GitHub a uživatelé si tak mohou nástroj snadno přizpůsobit a rozšířit podle svých specifických potřeb (Dutta, 2018).

Software nemusí být vhodný pro uživatele, kteří vyžadují vyvinutější anotační nástroje. VIA je desktopový nástroj, což znamená, že vyžaduje, aby si uživatelé stáhli a nainstalovali software na své lokální počítače (Dutta, 2018).

Souhrnně lze VGG Image Annotator považovat za dobrý anotační softwarový nástroj, který lze použít k anotaci velkých souborů dat obrázků pro použití v aplikacích strojového učení. Jeho snadné použití, přizpůsobivost a pokročilé funkce z něj činí cenný nástroj pro anotování a správu dat. Uživatelé by však opět měli pečlivě vyhodnotit své specifické potřeby a požadavky, než se rozhodnou používat VIA (Aljabri et al., 2022). Pracovní prostředí je na Obrázku 1.4 (robots.ox.ac.uk, 2023).



Obrázek 1.4: Pracovní prostředí softwaru VIA

1.5.4 LabelMe

LabelMe je online anotační nástroj, jenž je využíván v oblasti počítačového vidění. Dobrou vlastností softwaru je jeho přizpůsobivost. Podporuje řadu anotačních úloh, včetně ohraničujících boxů, polygonů a bodů, takže je vhodný pro různé případy použití. Poskytuje také možnost vytvářet vlastní anotační třídy a přizpůsobovat vzhled anotačního rozhraní podle preferencí uživatele (Russell, et al., 2008).

Software podporuje práci více uživatelů na stejném anotačním projektu současně. LabelMe navíc nabízí nástroje pro správu anotačních projektů, včetně možnosti sledovat průběh, spravovat uživatele a vytvářet zprávy (Russell, et al., 2008).

Jednou z výhod aplikace je její otevřenost. Nástroj je volně k dispozici k použití a jeho zdrojový kód je k dispozici na serveru GitHub. Uživatelé si tak mohou nástroj snadno přizpůsobit a rozšířit podle svých specifických potřeb. Vytvořené štítky lze ukládat jako soubory JSON přímo z aplikace. Repozitář LabelMe nabízí skript v jazyce Python, který vám pomůže převést anotace na formát PASCAL VOC (Datagen.tech, 2023).

Při používání tohoto softwaru je nutností se zamyslet nad některými potenciálními nevýhodami. Jednou z nich je příliš jednoduché pracovní prostředí viz Obrázek 1.5 (Ren, 2016). Dále je nástroj méně vhodný pro uživatele, kteří vyžadují pokročilejší anotační nástroje, jelikož nabízí jen hrstku základních funkcí. Kromě toho je LabelMe webový

nástroj, což znamená, že ke svému používání vyžaduje připojení k internetu. Formáty, jako jsou YOLO a COCO, nejsou podporovány. Je však možné stáhnout opensourcový skript, který umožní převést výstup LabelMe do formátu COCO (Torralba, et al., 2010).

LabelMe je obecně jednoduchý nástroj, který se používá v oblasti počítačového vidění. Možnost spolupráce a otevřený zdrojový kód z tohoto softwaru činí využívaný nástroj pro anotaci a správu dat (Torralba, et al., 2010).



Obrázek 1.5: Pracovní prostředí softwaru LabelMe

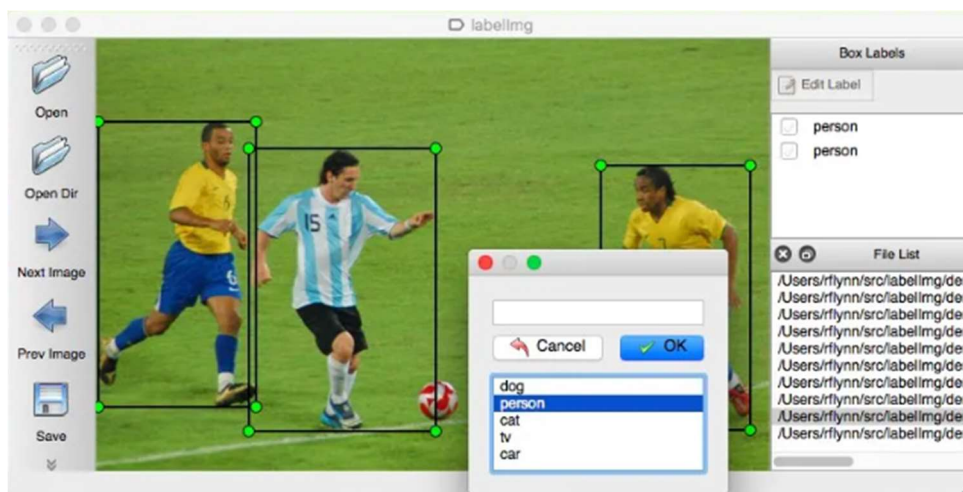
1.5.5 LabelImg

LabelImg je populární softwarový nástroj s otevřeným zdrojovým kódem, který se používá k anotování obrázků. Nástroj umožňuje uživatelům označovat objekty na obrázcích pomocí ohraničujících rámečků a poskytuje jednoduché rozhraní pro anotování dat. Nástroj se snadno používá a nabízí uživatelsky přívětivé rozhraní viz Obrázek 1.6 (Boesch, 2023), které usnadňuje rychlé anotování velkých souborů dat obrázků. Navíc poskytuje řadu možností přizpůsobení, včetně možnosti vytvářet vlastní anotační třídy a upravovat vzhled anotačního rozhraní. Podporuje širokou škálu formátů obrázků, včetně oblíbených formátů, jako jsou JPEG, PNG a BMP. Software rovněž podporuje označování v textovém formátu YOLO nebo ve formátu VOC XML. Uživatelům se doporučuje používat pro vytváření štítků výchozí formát VOC XML (Boesch, 2023).

Jelikož se jedná o formát sítě ImageNet, což je univerzální standard, tak je spojen s detekcí objektů, zatímco různé implementace YOLO mají různé formáty textových řádků. Navíc lze pomocí určitých nástrojů snadno převést formát VOC XML na jakýkoli jiný formát. Možností je například převést VOC XML do COCO JSON (Nelson, 2020).

Jednou z výhod LabelImg je jeho otevřenost. Nástroj je volně k dispozici k použití a jeho zdrojový kód je k dispozici na serveru GitHub. Uživatelé si tak mohou nástroj snadno přizpůsobit a rozšířit podle svých specifických potřeb (Nelson, 2020).

Nevýhodou tohoto softwaru může být menší přívětivost pro uživatele, kteří vyžadují propracovanější prostředí a nabídku nástrojů pro práci s obrázky, která je značně omezena. Kromě toho nástroj LabelImg vyžaduje, aby si uživatelé stáhli a nainstalovali software do svých lokálních počítačů (Nelson, 2020).



Obrázek 1.6: Pracovní prostředí softwaru LabelImg

1.5.6 Label Studio

Label Studio je softwarový nástroj s otevřeným zdrojovým kódem pro anotaci dat, který je široce využíván v oblasti strojového učení a umělé inteligence. Software je navržen tak, aby zjednodušil proces anotování velkých datových sad textu, obrázků a dalších typů dat a usnadnil uživatelům vytváření vysoce kvalitních trénovacích datových sad pro modely strojového učení (Gagny, 2023; Liubimov, 2020)

Label Studio funguje tak, že uživatelům umožňuje vytvářet anotační úlohy a nahrávat data pro anotaci. Uživatelé mohou vytvářet vlastní typy anotací a přizpůsobit si rozhraní podle svých preferencí. Po dokončení anotací může nástroj datasety exportovat v různých formátech vhodných pro již zmíněné použití v rámci strojového učení.

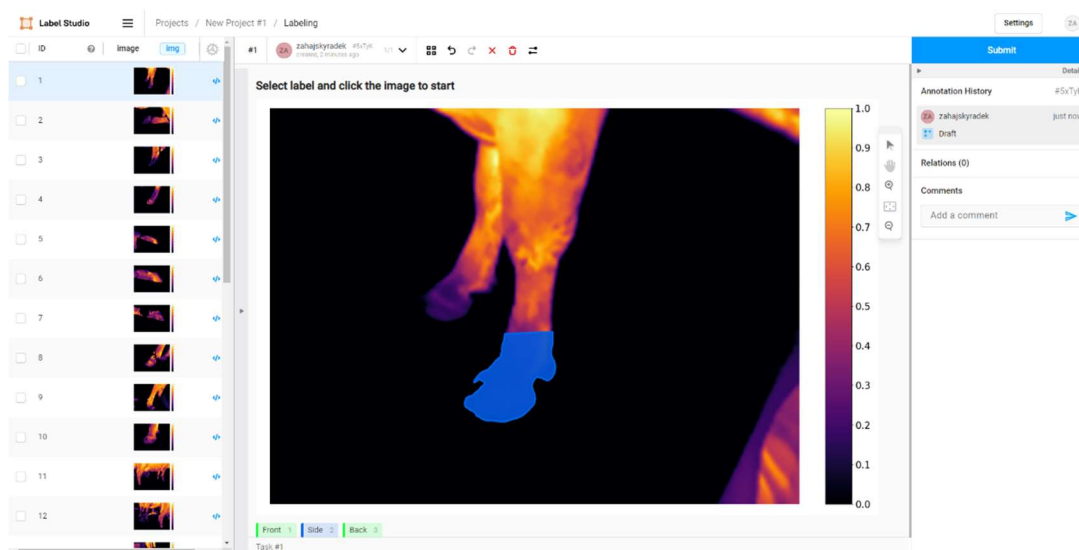
Podporuje širokou škálu typů anotací, včetně klasifikace textu, rozpoznávání pojmenovaných objektů, segmentace obrázků a detekce objektů (Liubimov, 2020).

Dobrou vlastností softwaru Label Studio je funkce spolupráce. Podporuje práci více uživatelů na stejném anotačním projektu současně, což přispívá ke zvýšení efektivity a přesnosti. Label Studio navíc poskytuje řadu nástrojů pro správu anotačních projektů, včetně sledování průběhu, spravování uživatelů a vytváření zpráv (Liubimov, 2020).

Velká výhoda Label Studia je jeho otevřenost. Nástroj je volně dostupný k použití a jeho zdrojový kód je k dispozici na serveru GitHub. Uživatelé si tak mohou nástroj snadno přizpůsobit a rozšířit podle svých potřeb. Dále zpracování pracovního prostředí, viz Obrázek 1.7 (autor), působí uspořádaně a propracovaně (Gagny, 2023).

Každý software je svým způsobem unikátní, a i Label studio disponuje nedostatky. Label Studio vyžaduje k nastavení a používání určité technické znalosti, což může být pro některé uživatele překážkou. Nicméně umožňuje jednotlivá nastavení a možnosti vytvořit pomocí programovacího jazyka, nebo ve zjednodušené formě již vytvořených nabídek pro úpravu. Určování pravidel může být z hlediska náročnosti pro nezkušené uživatele přítěží při práci s tímto softwarem (Exxact, 2022).

Při práci s obrázky je ovládání prostředí jednoduché, ale někdy nedostatečné. Před vytvořením projektu je nutno vybrat, jakou činnost s obrázky uživatel bude provádět a vybrat šablonu struktury prostředí, jelikož v nabídce nalezneme například šablonu pro práci s textem a zpracovávání audia. V této tématice je řešeno pouze anotování snímků, proto se vybírá tato možnost, která je však velice omezena (Exxact, 2022).



Obrázek 1.7: Pracovní prostředí softwaru Label Studio

1.6 Anotační formáty a datasety

Formát anotací pro obrázky je způsob zobrazování a ukládání informací o objektech nebo částech význačného prostoru na obrázku. Existuje několik různých formátů anotací, které umožní zaznamenat například bounding boxy ve formě obdélníku kolem každého objektu nebo oblasti výběru v obrázku. Souřadnice levého horního a pravého dolního rohu obdélníku jsou zaznamenány spolu s označením objektu či polygonální masky, kdy polygon zahrnuje nakreslení mnohoúhelníku kolem požadovaného objektu na snímku. Souřadnice vrcholů mnohoúhelníku se zaznamenají spolu se samotným výběrem (Pokhrel, 2020).

Výše zmíněné typy anotací se obvykle používají k trénování modelů strojového učení pro úlohy, jako je detekce objektů a segmentace. Každý anotační formát má svoji základní podobu. Jedná se o způsob zapisování dat nezávisle na počítačové platformě, jenž je určen pro přenos dat (Karatas, 2021).

1.6.1 MS COCO

MS COCO (Microsoft Common Objects in Context) je rozsáhlá datová sada pro rozpoznávání, segmentaci a popisování obrázků, která se skládá z více než 330 000 obrázků, z nichž každý je opatřen popisky objektů, segmentačními maskami a popisky. Účelem této datové sady je pokročit ve výzkumu počítačového vidění a umělé inteligence, zejména v oblastech rozpoznávání obrazu, detekce objektů a tvorby popisků k obrazům. MS COCO se široce používá jako srovnávací datová sada pro hodnocení výkonnosti modelů a algoritmů počítačového vidění. Tento formát se reprezentuje jako soubor JSON, který je široce využíván mezi anotačními softwary (Datagen.tech, 2023; Shah, 2023; Petrosyan, 2021).

JSON (JavaScript Object Notation) je odlehčený textový formát pro výměnu dat, jenž je snadno čitelný a zapisovatelný pro člověka a snadno analyzovatelný a generovatelný pro stroj. JSON je často využíván pro přenos dat mezi serverem a webovou aplikací jako alternativa k XML (Tyson, 2022).

V JSONu jsou data reprezentována dvojicí klíč-hodnota, podobně jako jsou data uložena ve slovníku v jazyce Python. Každý klíč je řetězec a hodnotou může být řetězec, číslo, pole nebo jiný objekt (Tyson, 2022).

Základní datová struktura viz Obrázek 1.8 (Cocodataset.org, 2023) obsahuje informace týkající se datasetu – rok vytvoření, verze a autora. Součástí je dále seznam obsahující informace o jednotlivých obrazech obsahující anotovaná data (Cocodataset.org, 2023).

První oddíl definuje čtyři základní subjekty, jako je primární informace, obrázek, anotaci a licence (Cocodataset.org, 2023).

Rozdělení náležitých informací je do těchto skupin – obsáhlejší informace, údaje obrázku a licenční podrobnosti (Cocodataset.org, 2023).

(červeně vyznačené atributy je potřeba vyplnit)

```
{
  "info"           : info,
  "images"        : [image],
  "annotations"   : [annotation],
  "licenses"      : [license],
}

info{
  "year"          : int,
  "version"       : str,
  "description"   : str,
  "contributor"   : str,
  "url"           : str,
  "date_created" : datetime,
}

image{
  "id"            : int,
  "width"         : int,
  "height"        : int,
  "file_name"     : str,
  "license"       : int,
  "flickr_url"    : str,
  "coco_url"      : str,
  "date_captured" : datetime,
}

license{
  "id"            : int,
  "name"          : str,
  "url"           : str,
}
```

Obrázek 1.8: Základní datová struktura MS COCO

Příklad datové struktury pro anotaci viz Obrázek 1.9 (Immersivelimit.com, 2019).

```
"annotations": [  
  {  
    "segmentation": [[510.66,423.01,511.72,420.03,...,510.45,423.01]],  
    "area": 702.1057499999998,  
    "iscrowd": 0,  
    "image_id": 289343,  
    "bbox": [473.07,395.93,38.65,28.67],  
    "category_id": 18,  
    "id": 1768  
  },
```

Obrázek 1.9: Struktura anotace MS COCO

Annotations spouští strukturu, ve které se přidělují jednotlivé hodnoty datům.

- *segmentation* určuje polygonový výběr se souřadnicemi bodů (x,y pozice pixelu),
- *area* vyznačuje oblast pixelů. V příkladném obrázku znázorněna velikost přibližně 702 pixelů, což je pro představu velice malá oblast,
- *iscrowd* má pouze hodnoty 1 nebo 0. Jednička nám dává informaci, že se jedná o více objektů, nula naopak znamená pouze jeden výběr,
- *image_id* koresponduje s hodnotou anotovaného obrázku,
- *bbox* poskytuje informace o ohraničujícím poli,
- *category_id* se přiřazuje se štítkem a určuje, co je ve výběru za objekt,
- *id* je samotná identita anotace (Immersivelimit.com, 2019).

Struktura musí být ukončena hranatou závorkou. Jedině takto se rozumí konci pole, které určuje hodnoty. Skutečná struktura již ve finální vyexportované podobě viz Obrázek 1.10 (Immersivelimit.com, 2019).

```
{ "id": 1411138, "category_id": 22, "iscrowd": 0, "segmentation": [[72.59, 177.35, 73.33, 16  
9.12, 77.08, 160.89, 80.07, 156.4, 85.31, 149.66, 88.3, 145.17, 92.04, 139.19, 96.53, 133.95, 100.2  
7, 127.96, 104.02, 125.72, 109.25, 124.97, 118.98, 124.22, 130.21, 122.72, 136.19, 122.72, 149.66,  
121.98, 161.64, 123.47, 169.87, 127.21, 170.62, 130.96, 167.62, 133.95, 163.13, 135.44, 158.64, 13  
4.7, 146.67, 140.68, 143.68, 148.91, 143.68, 155.65, 141.43, 160.14, 141.43, 164.63, 141.43, 165.3  
8, 136.19, 165.38, 133.2, 163.13, 131.7, 169.87, 132.45, 176.6, 132.45, 178.1, 132.45, 178.85, 127.  
96, 178.85, 122.72, 165.38, 122.72, 163.88, 120.48, 168.37, 116.74, 173.61, 114.49, 178.1, 110.75,  
177.35, 110.75, 171.36, 115.99, 161.64, 115.99, 153.4, 114.49, 151.16, 111.5, 149.66, 109.25, 151.9  
1, 102.52, 152.66, 98.78, 152.66, 96.53, 151.16, 91.29, 157.15, 83.06, 161.64, 77.82, 167.62, 76.3  
3, 170.62, 75.58, 173.61, 77.08, 176.6, 75.58, 177.35]], "image_id": 245915, "area": 2469.098400000  
0003, "bbox": [72.59, 121.98, 98.03, 56.87]]},
```

Obrázek 1.10: Struktura anotace s větším počtem bodů MS COCO

1.6.2 PASCAL VOC

PASCAL VOC (Visual Object Classes) je benchmarkový dataset a soutěž v oblasti počítačového vidění pro detekci, segmentaci a klasifikaci objektů. Skládá se z rozsáhlé sbírky anotovaných obrázků a sady standardních hodnotících protokolů pro testování výkonnosti modelů a algoritmů počítačového vidění (Everingham, a Winn, 2012).

Tato datová sada byla poprvé představena v roce 2005 a od té doby byla několikrát aktualizována a rozšířena. Obsahuje snímky dvaceti různých kategorií objektů, včetně lidí, zvířat, vozidel a objektů v interiéru. Každý obrázek je opatřen anotací s ohraničujícími boxy objektů, segmentací a štítky tříd (Everingham, a Winn, 2012).

Soutěž PASCAL VOC se stala jedním z nevlivnějších srovnávacích testů v oblasti počítačového vidění a na této sadě dat bylo vyhodnoceno a porovnáno mnoho nejmodernějších algoritmů pro detekci, segmentaci a klasifikaci objektů. Soutěže také pomohly podpořit pokrok v oboru a vedly k vývoji nových technik a algoritmů pro rozpoznávání objektů a porozumění prostředí (Khandelwal, 2019).

Mezi hlavní rozdíly formátů COCO a PASCAL VOC patří:

- Pascal VOC je založen na XML, na rozdíl od COCO, který má soubor JSON,
- V Pascal VOC vytváříme soubor pro každý obrázek v datové sadě. Na druhou stranu COCO vždy obsahuje jeden soubor pro celou sadu dat pro trénování, testování a validaci,
- Ohraničující rámeček v datových formátech Pascal VOC a COCO se liší (Khandelwal, 2019).

XML (Extensible Markup Language) je obecný značkovací jazyk, který byl vyvinut a standardizován konsorciem W3C (mezinárodní konsorcium, jehož členové společně s veřejností vyvíjejí webové standardy pro World Wide Web). Jedná se o zjednodušenou podobou staršího jazyka SGML (univerzální značkovací metajazyk). Umožňuje snadné vytváření konkrétních značkovacích jazyků pro různé účely a různé typy dat (Sciendo.com, 2022).

V XML jsou data reprezentována jako stromová struktura elementů, kde každý vlastní název, atributy a hodnotu. Elementy mohou mít podřízené elementy, což umožňuje reprezentovat složité datové struktury. Tento formát se používá pro sekvenci dat, v čemž soupeří s JSON (Sciendo.com, 2022).

Také je vhodný pro tvůrce webových stránek, jelikož jazyk XML je široce využíván při tvorbě webových stránek a databázových systémů (Khandelwal, 2019).

Jeho nepochybnou výhodou je jednoduchá úprava, srozumitelnost a podobná struktura s jazykem HTML (Hypertext Markup Language). Člověk velice rychle pochopí styl psaní a intuitivnost elementů přispívá k nabytí vědomostí tohoto jazyka. Další výhodou je jeho mezinárodní podpora, nepochybně má zastoupení mezi nejpulárnějšími programátorskými jazyky (Sciendo.com, 2022). Struktura vyobrazena na Obrázku 1.11 (Khandelwal, 2019).

```
<annotation>
  <folder>Kangaroo</folder>
  <filename>00001.jpg</filename>
  <path>./Kangaroo/stock-12.jpg</path>
  <source>
    <database>Kangaroo</database>
  </source>
  <size>
    <width>450</width>
    <height>319</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>kangaroo</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>233</xmin>
      <ymin>89</ymin>
      <xmax>386</xmax>
      <ymax>262</ymax>
    </bndbox>
  </object>
</annotation>
```

Obrázek 1.11: Struktura Pascal VOC

Níže jsou vysvětleny některé klíčové značky pro Pascal VOC.

- <Folder> – Složka, která obsahuje obrázky,
- <Filename> – Název fyzického souboru, který existuje ve složce,
- <Path> – Cesta soubory k obrázku,
- <Size> – Velikost obrázku na šířku, výšku a hloubku. Pokud je obrázek černobílý, bude hloubka 1. V případě barevných obrázků bude hloubka 3,
- <Object> – Podrobnosti o objektu. Pokud máte více anotací, pak se značka objektu s jeho obsahem opakuje.

Součásti značek objektů jsou:

-
1. <Name> – název,
 2. <Pose> – postoj,
 3. <Truncated> – zkrácený,
 4. <Difficult> – obtížnost,
 5. <Bndbox> – ohraničení,

name – Název objektu, který se snažíme identifikovat,

truncated – Tento výraz znamená, že ohraničující rámeček zadaný pro objekt neodpovídá celému rozsahu objektu. Pokud je objekt na obrázku viditelný částečně, nastavíme hodnotu truncated na 1. Pokud je objekt viditelný celý, nastavíme hodnotu truncated na 0,

difficulty – Objekt je označen jako obtížný, pokud je považován za obtížně rozpoznatelný. Pokud je objekt obtížně rozpoznatelný, pak nastavíme hodnotu obtížné na 1, jinak nastavíme hodnotu 0,

bndbox – Osově zarovnaný obdélník určující rozsah objektu viditelného na snímku (Khandelwal, 2019).

1.6.3 YOLO

You Only Look Once je oblíbený algoritmus detekce objektů, jehož architekturu v první verzi vytvořili Joseph Redmon, Santosh Divvala, Ross Girshick a Ali Farhadi. Tato neuronová síť se stala populárním díky své vysoké přesnosti, výkonu v reálném čase a jednoduchosti. Stejný název nese i formát vstupních anotací pro tuto síť. (Kundu, 2023; Jiang, et al., 2022; Nelson, 2021).

YOLO je specifický formát souboru používaný k ukládání anotací pro úlohy detekce objektů. Formát je navržen pro práci s vlastním algoritmem, ale lze jej použít i s jinými algoritmy detekce objektů. Strukturou je textový soubor, který obsahuje informace o ohraničujících polích a označeních tříd objektů v obraze (Kundu, 2023; Sharma, 2022).

Struktura je nenáročná a každý řádek souboru odpovídá objektu v obrázku. Obsahuje následující informace:

<object-class> <x> <y> <width> <height>

- <object-class> je označení třídy objektu,
- <x> a <y> jsou souřadnice středu ohraničujícího rámečku objektu,

-
- `<width>` a `<height>` jsou šířka a výška ohraničujícího rámečku.

V současné době existuje několik verzí algoritmu YOLO, každá s jinými vylepšeními a funkcemi (Sharma, 2022).

1.7 Veřejně dostupné datasety

Datová sada je soubor dat vytvořený pro projekt strojového učení. Sada obrazových dat zahrnuje digitální obrazy vytvořené pro testování, trénování a vyhodnocování výkonnosti algoritmů strojového učení a umělé inteligence, obvykle algoritmů počítačového vidění. Sady obrazových dat, jako je ImageNet, pomáhají algoritmům učit se identifikovat a rozpoznávat informace v obrazech a provádět související rozeznávací činnosti. Algoritmy umělé inteligence lze například vycvičit k označování fotografií, čtení poznávacích značek automobilů a identifikaci nádorů na lékařských snímcích. Algoritmy počítačového vidění mohou také transformovat obrazy nebo vytvářet zcela nové obrazy, což má řadu praktických aplikací (Devopedia, 2021; Brownlee, 2019).

Veřejně dostupným datasetem se rozumí volně přístupnou knihovnou snímků pro práci s obrazovými daty. Tato databáze všelijakých obrázků slouží jako dobrá základna pro strojové učení a rozdílem mezi druhy je v obsažených datech (Devopedia, 2021; Brownlee, 2019).

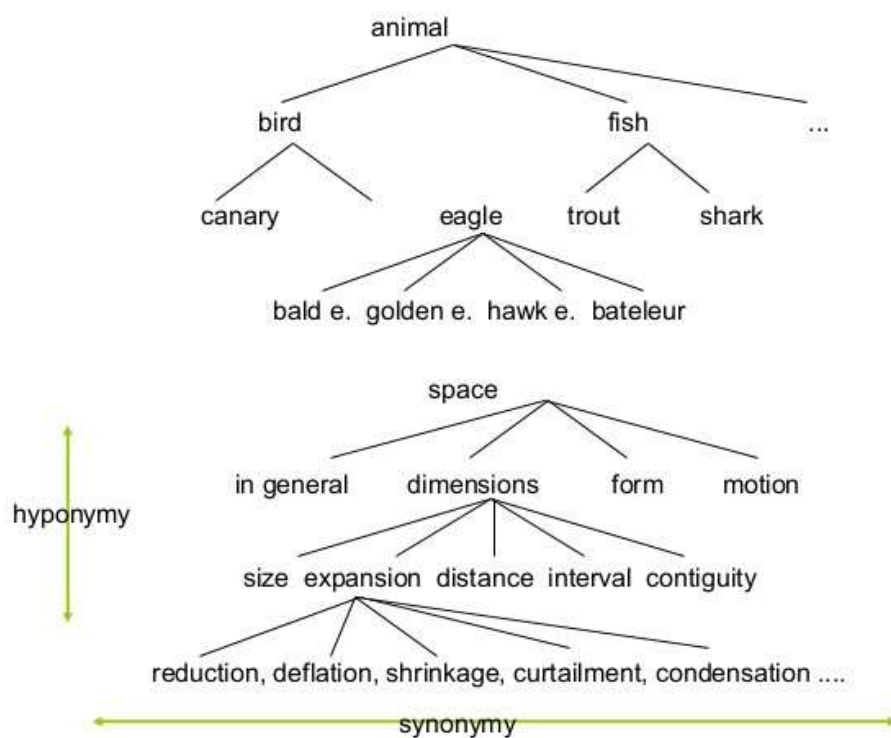
ImageNet je rozsáhlá databáze obrázků navržená pro použití ve výzkumu rozpoznávání vizuálních objektů. Vytvořili ji výzkumníci ze Stanfordovy univerzity a obsahuje více než 14 milionů obrázků, z nichž každý je označen sadou kategorií objektů. Je založen na myšlence, že výkonnost algoritmů počítačového vidění lze zlepšit jejich trénováním na velkém, různorodém souboru označených obrázků. Síť obsahuje více než 20 000 kategorií objektů, od běžných objektů až po specializovanější kategorie. Ve skutečnosti ImageNet tyto kategorie nedefinoval sám, ale odvodil je ze sítě WordNet (Devopedia, 2021; Brownlee, 2019), což je databáze anglických slov propojených významovými vztahy. Slova podobného významu jsou seskupena do souboru synonym, zjednodušeně nazývaného **synset**. Hypernymy jsou synsety, které jsou obecnější. Takto je "organismus" hypernymem slova "rostlina". Hyponymy jsou synsety, které jsou specifitější. Tedy "vodní" je hyponymum slova "rostlina" (Wordnet.princeton.edu, 2023).

Podstatná jména, slovesa, přídavná jména a příslovce jsou seskupena do souborů poznávacích synonym, z nichž každý vyjadřuje odlišný pojem. Výslednou síť významově příbuzných slov a pojmů lze procházet pomocí prohlížeče. WordNet je rovněž

volně a veřejně dostupný ke stažení. Díky své struktuře je užitečným nástrojem pro počítačovou lingvistiku a zpracování přirozeného jazyka (Wordnet.princeton.edu, 2023).

Existují důležité aspekty, které je dobré zmínit. Za prvé, WordNet propojuje nejen slovní tvary a řetězce písmen, ale i konkrétní významy slov. V důsledku toho jsou slova, která se v síti nacházejí v těsné blízkosti, významově rozlišena. Za druhé, označuje významové vztahy mezi slovy (Wordnet.princeton.edu, 2023).

Tato hierarchie viz Obrázek 1.12 (devopedia.org, 2020) je užitečná pro úlohy počítačového vidění. Pokud si model není jistý podkategorií, může jednoduše klasifikovat obrázek výše v hierarchii, kde je pravděpodobnost chyby menší. Pokud si například model není jistý, že se dívá na králíka, může ho jednoduše klasifikovat jako savce (Wordnet.princeton.edu, 2023).



Obrázek 1.12: Princip hierarchie ImageNetu

1.8 Barevné systémy

Barevné systémy, též označované jako prostory, představují způsob reprezentace barev pomocí číselných hodnot. Barevný systém definuje rozsah barev, které může zařízení, jako je monitor nebo tiskárna, zobrazit, a poskytuje metodu převodu mezi číselnými hodnotami a skutečnými barvami (Rosebrock, 2021).

V současné době se používá několik různých barevných systémů, z nichž každý má své výhody a nevýhody v závislosti na konkrétní aplikaci. Mezi oblíbené barevné systémy patří:

- RGB – běžně používaný pro digitální displeje, které zobrazují barvy kombinací červeného, zeleného a modrého světla,
- CMYK – využíván v tiskových médiích, které tisknou barvy pomocí kombinací azurového, purpurového, žlutého a černého inkoustu,
- HSV – barevný prostor, který rozděluje barevné informace na složky odstínu, sytosti a hodnoty, což může být užitečné při úpravách a manipulaci s obrázkem,
- LAB – tento barevný prostor je navržen tak, aby byl vnímán jednotněji, což znamená, že stejné vzdálenosti v barevném systému odpovídají přibližně stejným rozdílům ve vnímané barvě (Rosebrock, 2021).

Pochopení barevných systémů je důležité pro zajištění přesného zobrazení barev v různých zařízeních a médiích a pro dosažení konzistentních výsledků při digitálním zpracování obrazu a správě barev (Rosebrock, 2021).

Jedním z běžných použití barevných systémů v počítačovém vidění je rozpoznávání a segmentace objektů. Barevné systémy, jako jsou RGB a HSV, lze použít k extrakci barevných prvků z obrázků a k identifikaci objektů na základě jejich barevných charakteristik. V těchto aplikacích mohou být upřednostňovány barevné systémy jako LAB nebo HSL, protože umožňují intuitivnější a přesnější ovládání barevných hodnot (Rosebrock, 2021).

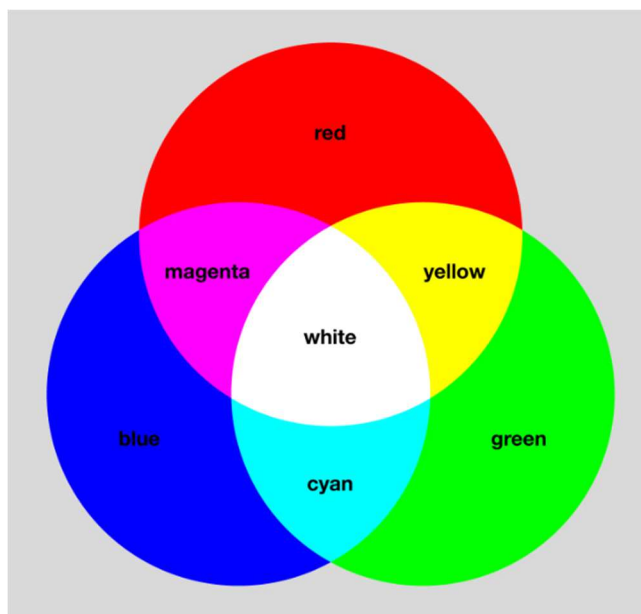
Konvoluční neuronové sítě (CNN), populární typ modelu hlubokého učení používaného v počítačovém vidění, často přijímají jako vstup surové hodnoty pixelů. Předzpracování obrázků převodem do jiného barevného systému, například LAB, a následnou úpravou však lze zvýšit přesnost a efektivitu modelu tím, že sníží množství potřebných vstupních dat a normalizuje hodnoty barev (Rosebrock, 2021).

1.8.1 RGB

V nejběžnějším barevném systému RGB jsou barvy reprezentovány červenou, zelenou a modrou složkou. Z technického hlediska popisuje RGB barvu jako trojici složek, které mohou nabývat hodnot v rozmezí 0 až 255, přičemž trojice (0, 0, 0) představuje černou a (255, 255, 255) bílou (Kovalčík, 2020).

RGB je příkladem přídavného barevného systému. Čím více každé barvy se přidá, tím je pixel jasnější a tím více se blíží bílé barvě viz Obrázek 1.13 (Zelazko, 2022). Problém tohoto systému spočívá v tom, že není nejvhodnějším barevným systémem pro vývoj aplikací založených na konvenčních metodách počítačového vidění. Ve skutečnosti se primárně používá k zobrazení barev na monitoru (Kovalčík, 2020).

Protože barva RGB je definována jako hodnota tří veličin s každou hodnotou v rozsahu od 0 do 255, můžeme si představit, že barevný systém RGB obsahuje součin $256 \times 256 \times 256$. Výsledkem je 16 777 216 možných barev v závislosti na tom, kolik červené, zelené a modré barvy umístíme do jednotlivých políček (Zelazko, 2022).



Obrázek 1.13: RGB prostor se spojením barev

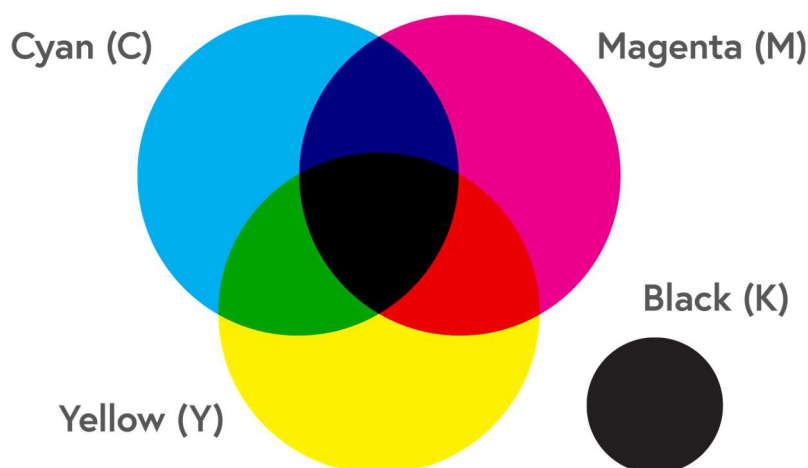
1.8.2 CMYK

CMYK je barevný systém, který se široce používá v tiskařském průmyslu k reprezentaci barev pro účely tisku. Jedná se o subtraktivní barevný model, což znamená, že začíná bílou barvou a přidává barvy postupným odebráním (pohlčováním) bílého světla (Andersen, 2020; Artuffo, 2019).

V CMYKu jsou barvy reprezentovány jako kombinace azurové, purpurové, žluté a černé barvy viz Obrázek 1.14 (mixam.com, 2023). Černý inkoust se často označuje jako "klíčová" barva, odtud název CMYK. Kombinací těchto čtyř barev lze získat širokou škálu barev, které jsou vhodné pro tisk (Andersen, 2020).

Každá barva v barevném systému CMYK je reprezentována procentuální hodnotou mezi 0 a 100, kde 0 představuje žádný inkoust a 100 maximální sytost inkoustu. Například barva reprezentovaná jako (50 %, 0 %, 75 %, 0 %) by se skládala z 50 % azurové, 0 % purpurové, 75 % žluté a 0 % černé barvy (Andersen, 2020).

Jednou z výhod CMYK oproti jiným barevným systémům je, že je speciálně navržen pro tisk, což znamená, že barvy lze přesně reprodukovat na široké škále tiskových zařízení. Protože se však jedná o subtraktivní barevný model, může být někdy obtížné přesně reprezentovat barvy, které jsou mimo paletu dostupných barev inkoustu (Artuffo, 2019).



Obrázek 1.14: CMYK prostor se skládáním

1.8.3 HSV

HSV (Hue, Saturation, Value) je systém barev, který reprezentuje barvy na základě tří složek: odstínu, sytosti a hodnoty. Často se používá v počítačové grafice, zpracování obrazu a počítačovém vidění, protože rozděluje barevné informace na intuitivní a vjemově smysluplné složky (Kang, et al., 2021).

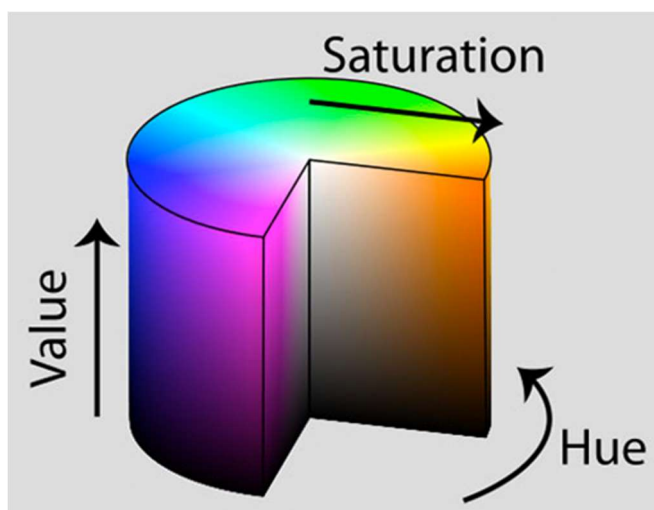
Odstín je barevnou složkou barevného systému a je obvykle reprezentován jako úhel kolem barevného kola, přičemž červená barva má hodnotu 0 stupňů, zelená 120 stupňů a modrá 240 stupňů. Ostatní barvy jsou reprezentovány jako úhly mezi těmito základními barvami (Kang, et al., 2021).

Sytost představuje intenzitu nebo čistotu barvy a je obvykle reprezentována jako procentuální hodnota mezi 0 % a 100 %, kde 0 % je zcela nenasycená barva (šedá) a 100 % je nejčistší, nejživější barva (Howlett, 2017).

Hodnota představuje jas nebo světlost barvy a je obvykle reprezentována jako procentuální hodnota mezi 0 % (černá) a 100 % (bílá) (Howlett, 2017).

Rozdělením barev na složky odstínu, sytosti a hodnoty umožňuje barevný systém HSV intuitivnější a přesnější ovládání hodnot barev. Například změnou složky odstínu lze upravit celkovou barvu obrázku, zatímco změnou složek sytosti a hodnoty lze upravit intenzitu a jas konkrétních barev v obrázku (Guha, 2021).

Celkově je barevný systém HSV výkonným nástrojem pro zpracování obrazu a aplikace počítačového vidění, který poskytuje flexibilní a přehledný rámec pro manipulaci s barevnými informacemi a jejich analýzu (Guha, 2021). Systém se směrem hodnot je vyobrazen na Obrázku 1.15 (Rosebrock, 2021).



Obrázek 1.15: HSV prostor se směrem hodnot

1.8.4 $L^*a^*b^*$

LAB je specifický systém barev, který byl navržen tak, aby byl vjemově jednotnější než jiné barevné systémy, jako je RGB nebo CMYK. To znamená, že stejné vzdálenosti v barevném systému odpovídají přibližně stejným rozdílům ve vnímané barvě (Rosebrock, 2021).

LAB je trojrozměrný barevný systém, jenž rozděluje barvy na tři složky: L^* (světlost), a^* (červeno-zelená osa) a b^* (modro-žlutá osa) viz Obrázek 1.16 (Rosebrock, 2021).

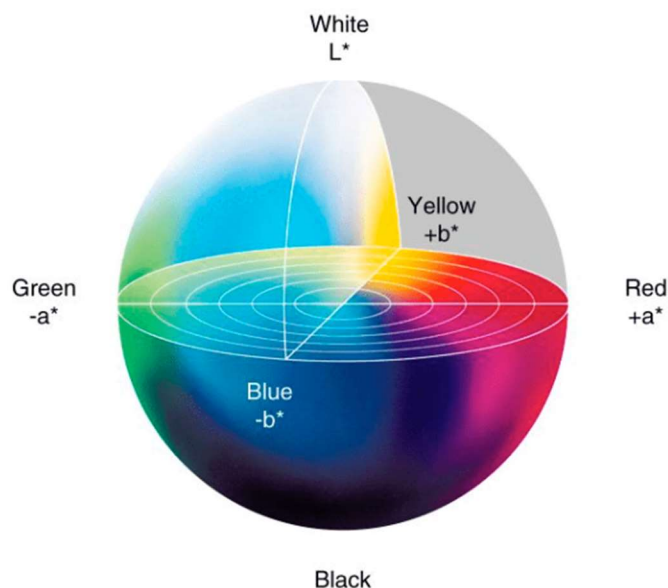
Složka L^* představuje světlost nebo tmavost barvy a je reprezentována jako hodnota mezi 0 a 100, kde 0 je černá a 100 je bílá (Rosebrock, 2021).

Parametry a^* a b^* představují informaci o barvě, přičemž kladné hodnoty a^* označují červenost a záporné hodnoty zelenost a kladné hodnoty b^* žlutost a záporné hodnoty modrost (Rosebrock, 2021).

LAB oproti jiným barevným systémům je navržen tak, aby byl nezávislý na zařízení, což znamená, že jej lze použít k přesnému vyjádření barev bez ohledu na konkrétní použité zařízení nebo médium (Rosebrock, 2021).

Tento systém se často používá v aplikacích správy barev a porovnávání barev a také v aplikacích zpracování obrazu a počítačového vidění, které vyžadují přesnou analýzu barev nebo manipulaci s nimi. Lze ji použít k úpravě vyvážení barev obrazu nebo ke sladění barev produktu v různých výrobních sériích (Sensing.konicaminolta.us, 2023).

Celkově je LAB výkonný a všestranný barevný systém, který se široce používá v řadě aplikací, kde je důležité přesné a konzistentní zobrazení barev (Sensing.konicaminolta.us, 2023).



Obrázek 1.16: $L^*a^*b^*$ prostor

1.8.5 Termální snímky

Termosnímek, známý také jako termální fotografie nebo termogram, je snímek, který zachycuje infračervené záření vyzařované objekty v prostředí. Termokamera nezachycuje viditelné světlo jako běžná fotografie, ale teplo vyzařované objekty ve formě infračerveného záření a na základě těchto informací vytváří obraz (Knížková, a Kunc, 2020).

Termokamera vychází ze skutečnosti, že všechny objekty vyzařují infračervené záření úměrné své teplotě. Teplejší objekty vyzařují více záření, zatímco chladnější objekty vyzařují méně. Termokamera používá k detekci tohoto záření senzory a vytváří obraz, jenž znázorňuje změny teploty v prostředí (Knížková, a Kunc, 2020).

Tento typ snímků se často používají v průmyslových, lékařských a vojenských aplikacích, kde mohou poskytnout důležité informace o změnách teploty, které nemusí být viditelné pouhým okem. Termosnímky lze využít k detekci tepelných ztrát v budovách, identifikaci horkých míst v elektrických zařízeních, lokalizaci netěsností v potrubí nebo dokonce k detekci horečky u lékařských pacientů (Knížková, a Kunc, 2020).

Tato technologie má nepochybné zastoupení i v zemědělství. Termovizní zobrazování má několik aplikací při monitorování zdravotního stavu hospodářských zvířat. Termovizní snímkování pomáhá zemědělcům odhalovat včasné příznaky zdravotních problémů u jejich hospodářských zvířat. Zjišťováním teplotních rozdílů na povrchu zvířat mohou zemědělci identifikovat oblasti zánětu nebo infekce, které nemusí být pouhým okem viditelné. Například kulhání je u hospodářských zvířat velkým problémem a může být obtížné ho včas odhalit. Termokamera napomáhá zemědělcům pomoci identifikovat oblasti zánětu nebo infekce, které mohou být příčinou kulhání zvířat. To může zemědělcům pomoci přijmout nápravná opatření dříve, než se stav zhorší. Dalším častým problémem je mastidita dojníc, která může vést ke snížení produkce mléka a v závažných případech dokonce k úhynu. Termokamera může zemědělcům pomoci odhalit včasné příznaky mastitidy tím, že opět zjistí teplotní rozdíly na povrchu vemene (Knížková, a Kunc, 2021).

Celkově je termovize cenným nástrojem pro monitorování zdravotního stavu hospodářských zvířat v zemědělství, protože může zemědělcům pomoci identifikovat potenciální zdravotní problémy dříve, než se stanou vážnějšími a ovlivní produktivitu jejich hospodářských zvířat (Knížková, a Kunc, 2021).

Termosnímky se liší na základě způsobu vyobrazení teplot. Takzvané palety mění vzhled scény a zvýrazní klíčové oblasti termálního snímku, aniž by se změnily teplotní údaje. Tepelné palety jsou do značné míry záležitostí preferencí uživatele, ale v různých prostředích nebo situacích může být výhodnější jedna paleta než druhá.

Je důležité také zmínit, že počet pixelů na termálním snímku je dán rozlišením kamery. Snímače s vyšším rozlišením generují snímky s vyšším počtem pixelů a obecně poskytují jasnější výsledky (Thermascan.co.uk, 2019).

Mezi termografy oblíbená paleta takzvaného Tekutého železa. Jedná se o univerzální paletu, jenž rychle identifikuje tepelné rozdíly pomocí barev pro zobrazení rozložení tepla a jemných detailů. Horké objekty jsou zobrazeny ve světlejších, teplých barvách, zatímco chladnější objekty v tmavých, studených barvách (Thermascan.co.uk, 2019).

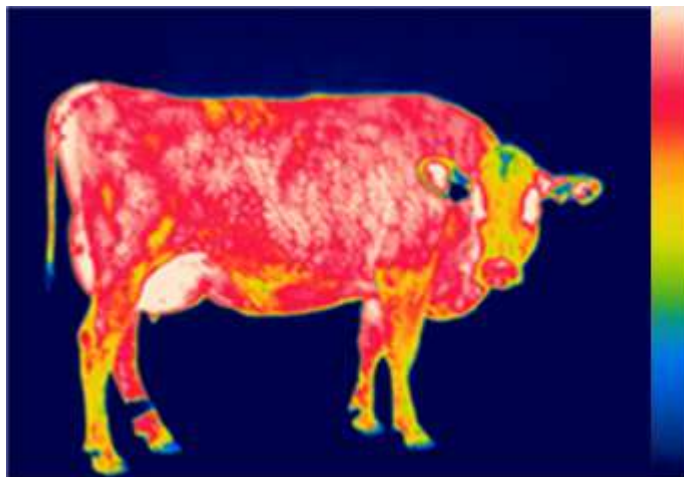
Tato paleta v programovacím jazyku Python je nazvána škálou Inferno (Hunter, et al., 2012). Termosnímek s barevnou škálou inferno viz Obrázek 1.17 (Knížková, a Kunc, 2021).



Obrázek 1.17: Ironbow termosnímek s inferno škálou

Rainbow obdobná paletě Ironbow, kdy teplé barvy představují nejteplejší části obrazu a studené barvy nejchladnější části, nicméně přidává do směsi více barev. Hodí se k přesnému určení objektů v prostředí s minimálními tepelnými rozdíly (Thermascan.co.uk, 2019). Tuto paletu v Pythonu nahrazuje formát Jet (Hunter, et al., 2012).

Dnes již existuje technologie Rainbow HC, jenž přidává do obrazu více kontrastu (High Contrast). Ve srovnání s jinými paletami si lze všimnout velice jemných teplotních přechodů (Thermascan.co.uk, 2019). Termální snímek v režimu rainbow se škálou jet viz Obrázek 1.18 (Knížková, a Kunc, 2021).



Obrázek 1.18: Rainbow snímek s jet škálou

Nejčastěji používaná paleta ve všech oblastech je White Hot, která zobrazuje teplejší objekty bílou barvou a chladnější objekty černou barvou. Palety ve stupních šedi nabízejí jednoduchost pro scény s širokým teplotním rozpětím a generují snímky s realistickými detaily (Thermascan.co.uk, 2019).

Univerzálnost palety White Hot ji činí atraktivní pro použití v proměnlivých krajinách a městských oblastech (Thermascan.co.uk, 2019). Termosnímek s paletou viz Obrázek 1.19 (Knížková, a Kunc, 2021).



Obrázek 1.19: White hot termosnímek se škálou

2 Cíl práce

Cílem bakalářské práce je vytvoření kvalitního datasetu, který je základem pro systémy počítačového vidění ke zpracování a analýze termálních snímků paznehtů skotu a rozpracování problematiky využití termografických snímků v živočišné výrobě včetně stručného popisu možnosti automatického zpracování obrazu pomocí metod umělé inteligence.

3 Metodika

Hlavním tématem této části je popis přípravy datasetů a jejich samotné zpracování. Jedná se o nejdůležitější problematiku, jelikož je za potřebí velké množství kvalitně anotovaných dat pro dosažení požadovaných výsledků.

3.1 Vstupní data

Základním materiálem pro práci na datasetech v bakalářské práci jsou termosnímky získané pomocí kamer FLIR E6 a FLIR E96. U kamery FLIR E6 je důležité před pořízením snímku nastavit emisivitu. Verze E96 vyžaduje předvolit emisivitu a vlhkost.

Farma Mokřý Lom, spadající pod zemědělské družstvo Ločnice, ve spolupráci s Fakultou zemědělskou a technologickou poskytla možnost získání snímků v běžném provozu. V rámci snímkování nebyla činnost pracovníků nijak výrazně omezena.

Samotné snímky slouží jako zdroj pro zpracování a učení umělé inteligence pro potřeby výzkumu Fakulty zemědělské a technologické Jihočeské univerzity v Českých Budějovicích. Pořízeno bylo 1609 termosnímků s velkou diverzitou objektů na snímcích, jelikož je potřeba různorodost a obsáhlá databáze pro učení CNN. V každém obrázku se nachází alespoň jeden pazneht, v některých i více, aby bylo možné rozpoznat i více objektů v jednom záběru. Termografické snímky byly převedeny do formátu JPG. Příklad klasického snímku na Obrázku 3.1 (autor).



Obrázek 3.1: Klasický snímek paznehtu

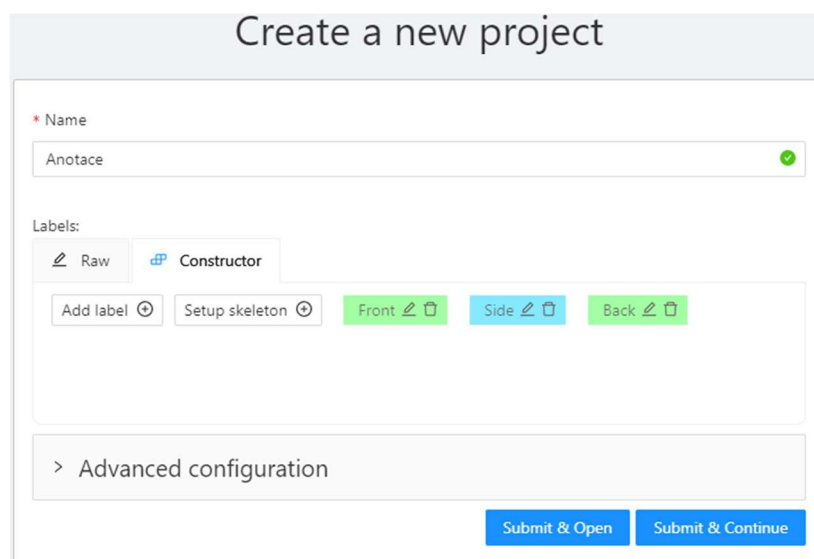
3.2 Příprava dat

Pro práci s daty byl použit anotační software CVAT. Důvodem ručního anotování byl požadavek vymezit pouze část paznehtu, přičemž se v obrazu často nachází celá končetina či další tepelně podobné objekty. Konvenční metody zpracování obrazu by tento požadavek splnili jen stěží, nebo vůbec. Je vhodné je aplikovat v případě, kdy je třeba anotovat celý objekt, který je od dalších objektů v obrazu vizuálně oddělen.

Prvním krokem je vytvoření hlavního projektu pomocí nabídky vytvořit nový projekt viz Obrázek 3.2 (autor). Pro jednoduchost je zvolen název Anotace.

V nabídce *Labels* pod možností *Constructor* se vytvořily 3 štítky s názvy *Front*, *Side*, *Back*. Přední a zadní mají barvu zelenou a boční modrou. *Constructor* je nástroj pro jednoduché vytvoření již zmíněných štítků. Nabídka *Raw* umožňuje ručně napsat script programátorským jazykem.

Pokročilá nastavení neměníme a nastavení potvrdíme jedním z tlačítek *submit*.



The screenshot shows a web form titled "Create a new project". At the top, there is a text input field for "Name" with the value "Anotace" and a green checkmark icon. Below this, there is a section labeled "Labels:" with two tabs: "Raw" and "Constructor". The "Constructor" tab is active. Underneath the tabs, there are three buttons: "Add label" (with a plus icon), "Setup skeleton" (with a plus icon), and three colored buttons: "Front" (green), "Side" (blue), and "Back" (green). At the bottom of the form, there is a section labeled "> Advanced configuration" and two blue buttons: "Submit & Open" and "Submit & Continue".

Obrázek 3.2: Nabídka vytvoření nového projektu

Dále je nutno dataset nahrát do knihovny úkolů v již vytvořeném projektu. Tuto akci umožňuje nabídka pro vytvoření nového úkolu viz Obrázek 3.3 (autor). Název zvolíme stejný a vybereme možnost právě požadovaného projektu.

Podskupinu zde nemusí brát v potaz, jelikož se jedná o jednoduchou strukturu skladby. Štítky budou k úkolu automaticky přiřazeny a vybrány díky volbě projektu, pod který tento úkol spadá.

Následně se vyberou soubory, které chceme nahrát do knihovny projektu. Možnost vybrání souborů nám umožňuje nahrát data přímo z uložiště zařízení.

Pokročilým nastavením není potřeba se zabývat, jelikož se obrázky nahrají ve formátu nejvhodnějším pro CVAT, tedy CVAT for images. Kvalita obrazu je nastavena na základních 70 %. V této hodnotě není komprese hodnot snímků příliš znatelná, ztráta detailů je patrna jen při velkém zvětšení.

Posledním krokem je již samotné potvrzení nabídky tlačítky *submit*. První nás vezme zpět do nabídky úkolů, kde již nalezneme námi vytvořený dataset. Druhé přímo otevře nahraný soubor obrázků a spustí se samotné prostředí pro práci s jednotlivými snímky.

Create a new task

Basic configuration

* Name
Anotace

Project
Anotace

Subset
Input subset

Labels
Project labels will be used

* Select files
My computer Connected file share Remote sources Cloud Storage

Click or drag files to this area
You can upload an archive with images, a video, or multiple images

382 files selected

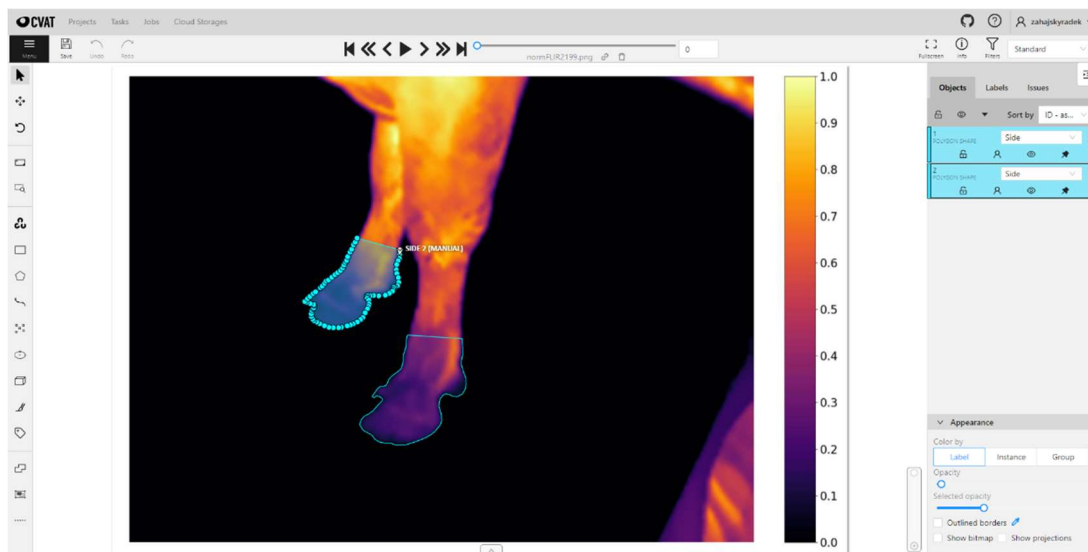
> Advanced configuration

Submit & Open Submit & Continue

Obrázek 3.3: Nabídka vytvoření nového úkolu

3.3 Pracovní prostředí

Nejdůležitější prostředí anotačního softwaru CVAT je pracovní plocha viz Obrázek 3.4 (autor). Zde jsou prováděny veškeré potřebné práce pro anotaci a úpravu na snímků. V levé horní části plochy se nachází hlavní nabídka, která obsahuje možnosti *importu*, *exportu*, *odstranění*, *otevření* a další funkce. Tyto jsou však ty nejdůležitější. Napravo od nabídky menu je možnost práci uložit, vrátit krok zpět a opětovný krok dopředu, což jsou velice užitečné akce, které umožňují procházet kroky při práci.








Obrázek 3.4: Pracovní plocha v softwaru CVAT

Uprostřed horní lišty se nachází panel listování mezi snímky. Nabízí možnost skočit na začátek nebo konec projektu, posunout o deset snímků nebo přejít na následující či předchozí. Tlačítko *play* pustí plynulý posun celého projektu, tuto možnost však není potřeba využívat. Napravo od tlačítek se nachází posuvník s číslem aktuálního snímku pro možnost rychlého posunutí na požadovaný snímek. V pravé části horní lišty je možnost přepnutí do celé obrazovky, info pro anotační statistiku a filtry. Tyto funkce není potřeba využívat.

V levé části plochy se nachází panel nástrojů. Jedná se o hlavní nabídku potřebných funkcí, které interagují se samotnými snímky. *Kurzor* je výchozí funkce, která nám umožňuje pohybovat se snímkem na pracovní ploše a vybírat různé objekty, vytvořené při práci na obrázku. Druhá funkce v pořadí shora dolů je *posun* snímku. Následuje možnost *otočení* o 90° doleva nebo doprava. Tyto funkce jsou určeny pro manipulaci se snímkem a nijak nezasahují do samotného anotování. Další dvě možnosti pracují se zvětšením obrazu, což je poměrně důležitá funkce, jelikož je zapotřebí zvětšení obrázku k přesné anotaci okrajů. Nicméně tyto funkce jednoduše nahrazuje

zvětšování pomocí kolečka myši. Člověk intuitivně kolečko používá buď k posouvání obrazu například na internetovém prohlížeči, nebo k již zmíněnému zvětšování obrazů. Vše viz Tabulka 1.










Tabulka 1: Nástroje softwaru CVAT

Nástroj	Popis
	Kurzor
	Posun snímku
	Rotace snímku
	Přizpůsobení snímku velikosti pracovní plochy
	Výběr regionu pro zvětšení

Následující funkce slouží pro anotaci snímků a jsou znázorněny v Tabulce 2.




OpenCV využívá umělé inteligence k rozpoznání hran na základě předvytvořené databáze. *Obdélník* je jedním z nejjednodušších formátů anotování. Je dán dvěma body, což vytváří nenáročná data. Výběrem je celý obsah obdélníku. *Polygon* umožňuje prakticky obkreslit jakýkoliv objekt. Každým kliknutím se určí jedna souřadnice a následně se prostor mezi body spojí příčnou čarou. Výběr je opět vnitřní obsah. Práce s touto funkcí je vysvětlena v kapitole 3.4. *Křivka* pouze spojuje vytvořené body, tudíž není vhodná pro výběr objektů. Nástroj je spíše určen pro vytváření čar. *Body* vytváří pouze body na obrazu, nikoliv spojovací čáry mezi nimi. S pomocí funkce *elipsa* vytvoříme jakýkoliv tvar elipsy, který je potřeba. Funkce *krychle* už pracuje na bázi třech dimenzí v obrazu. Vytvoří se krychle o jakékoliv velikosti a funkce automaticky vygeneruje protější stranu na souřadnici hloubky. Tento nástroj má velký potenciál, jelikož otevírá nové možnosti výběru objektu ze snímku. V současné době systém CVAT nepodporuje vyřezávání průhledných otvorů v polygonech pomocí *mask*. Vytvoření *značky* nám umožňuje přiřadit jeden ze štítků objektu. Použitím nástrojů pro tvorbu anotací není nijak zasahováno do původního obrazu, jedná se o vizualizaci přidávaných bodů v nové vrstvě.

Tabulka 2: Nástroje softwaru CVAT

Nástroj	Popis
	OpenCV
	Vytvořit nový obdélník
	Vytvořit nový polygon
	Vytvořit novou křivku
	Vytvořit nové body
	Vytvořit novou elipsu
	Vytvořit novou krychli
	Vytvořit novou masku
	Vytvořit značku

Poslední tři funkce slouží k úpravě vytvořených polygonů viz Tabulka 3. První slučuje prolínající se anotace, druhá slučuje anotace do jedné skupiny výběru a poslední nástroj polygony rozděluje podle potřeby.

Tabulka 3: Nástroje softwaru CVAT

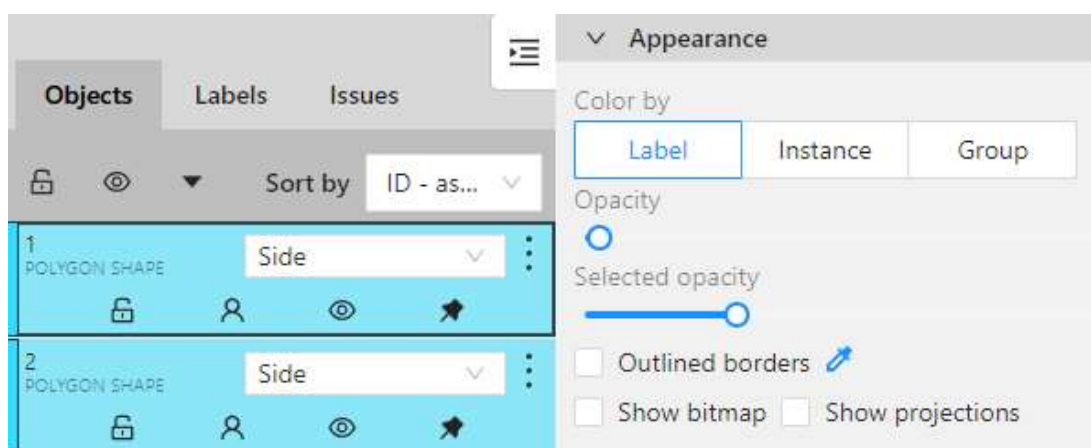
Nástroj	Popis
	Slučování prolínání
	Slučování výběrů
	Rozdělení výběrů

Na pravé straně pracovního prostředí viz Obrázek 3.5 vlevo (autor), se nachází nabídka, která zobrazuje všechny vytvořené výběry. Každému určuje číslo, jaký je použitý tvar, barevné rozdělení se svým štítkem, zda se jedná o objekt z pohledu předního,

bočního nebo zadního. Tento štítek lze po vytvoření výběru změnit, což je velice užitečná funkce v případě špatného výběru při vytváření. Každý z výběrů navíc nabízí další akce. Zamknout výběr, přepnout možnosti zakrytí, skrýt nebo zobrazit a připnout. Lze také vytvořit URL objekt (Uniform Resource Locator), udělat kopii, upravit, propagovat na další snímek, změnit orientaci, vložit do pozadí nebo popředí a smazat.

V druhé záložce se provádí akce s celými štítky. Určuje se priorita, zamknutí a zobrazení výběrů s přiřazenými štítky. Třetí záložka slouží pro kontrolu chyb. Nabídka na spodní části panelu viz Obrázek 3.5 vpravo (autor), už upravuje jen vyobrazení snímku a výběru a nemá vliv na anotování.

Pracovní plocha uprostřed prostředí obsahuje snímek, na kterém se anotace provádí. Jednotlivé body vytvořených anotací mají pevné souřadnice a tím pádem nedochází k rozchodu anotačního výběru při oddálení nebo přiblížení snímku.



Obrázek 3.5: Nabídky pravého panelu

3.4 Anotace dat

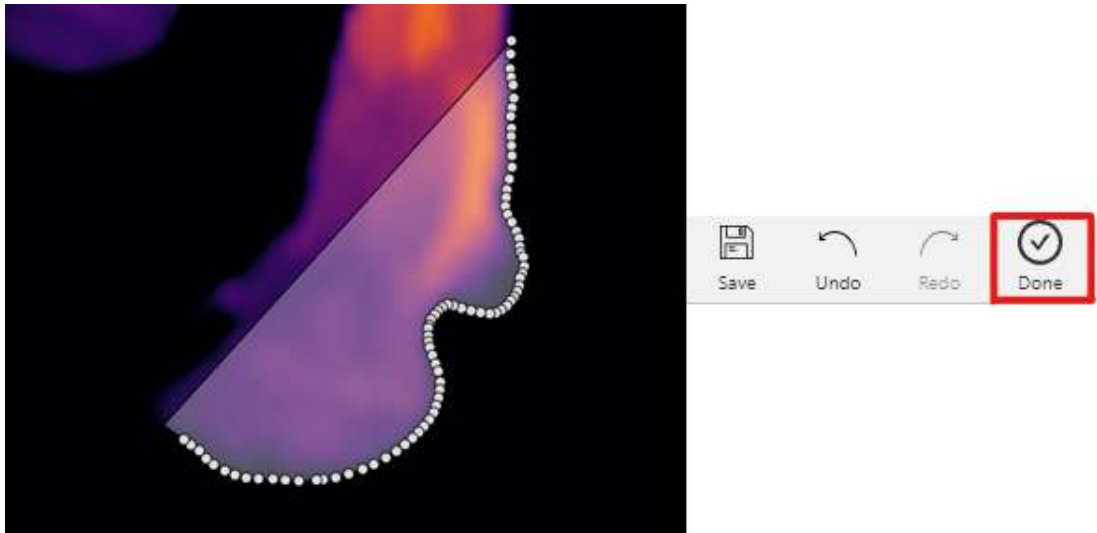
Při práci se snímky se využíval pouze jeden nástroj, a to anotovat pomocí polygonu. Umožňuje určit neomezený počet bodů, podle kterých se následný výběr po rovných spojovacích přímkách vytvoří. Tato funkce je pro ruční anotování velice důležitá, jelikož je zapotřebí kvalitních anotací s přesným výběrem okrajů.

Pro vytvoření polygonového výběru klikneme na nástroj *Polygonu* a vybereme jeden ze tří nabízených štítků. Počet bodů nevyplňujeme, jelikož množství se určí podle samotného zadávání při anotování. Zpravidla platí, že čím více bodů, tím složitější anotace, a tím pádem jsou data přesnější, avšak reakce softwaru se s výrazně rostoucím počtem bodů zpomalují. Další možností je vytvořit polygon pomocí možnosti *Shape* a *Track*.

Možnost *Shape* vytvoří tvar, který se uloží k aktuálnímu snímku. Naproti tomu funkce *Track* zaznamenává stopu, která se ručně vytvoří. Tento výběr se identicky přenáší na další snímky podle výběru na pravé liště objektů. Usnadňuje to práci se stejnými snímky, nicméně pro anotaci různých snímků je tento nástroj nepoužitelný.

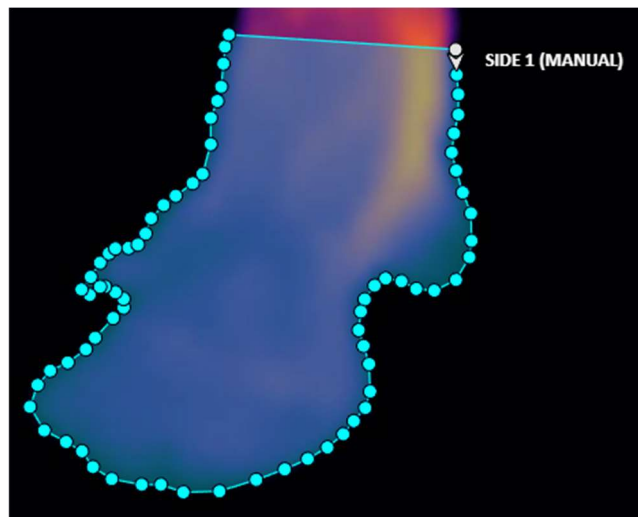
Začneme tedy postupně obkreslovat objekt, který chceme anotovat. Určujeme body klinutím levého tlačítka myši viz Obrázek 3.6 vlevo (autor). Při postupu je důležité klást důraz na přesnost a výběr musí být přesně na hranici objektu, který chceme anotovat. Pokud hrana objektu není ze snímku znatelná například kvůli focení z boku, kdy skot může mít nohy u sebe, je nutno pazneht orientačně vystihnout. Je potřeba zamezit prolínání více objektů do anotačního výběru, aby se výběr podobal tvaru paznehtu skotu.

Pro dokončení anotace se výběr potvrdí tlačítkem, jenž se objeví vedle možnosti vrátit krok dopředu v horní liště. Tímto ukončíme určování bodů a výběr se uzavře mezi prvním a posledním viz Obrázek 3.6 vpravo (autor).



Obrázek 3.6: Postup obkreslování vlevo, tlačítko pro potvrzení výběru vpravo

Výstupem je hotový výběr objektu, který se přiřadí k vybranému snímku. Takto je přidružen polygon kopírující hrany paznehtu v obrázku. Tímto způsobem vytvořený výběr je již finální forma dat a nadále se neupravuje. Rozpracovaný projekt je zapotřebí průběžně ukládat, jelikož CVAT nepodporuje automatické ukládání dat. Zhotovený výběr viz Obrázek 3.7 (autor).



Obrázek 3.7: Ukončený výběr

3.5 Export dat

Po dokončení celého datasetu je nutné data vyexportovat v určitém formátu, jenž je zapotřebí při dalším využití anotovaných snímků. Jedná se o klíčovou vlastnost přidruženou k obrázku a CVAT podporuje velké množství formátů. Díky tomu může být tento anotační software velice využíván. Například COCO, YOLO, PASCAL VOC a ImageNet.

Pro export dat anotací zvolíme v menu možnost Export job dataset viz Obrázek 3.8 (autor). Tato akce otevře novou nabídku viz Obrázek 3.9 (autor), kde již máme možnost vybrat formát, ve kterém dataset vyexportujeme. Výběr je tvořen z 24 formátů, které CVAT podporuje. V této práci je využit formát COCO (MSCOCO).

Důležité je vybrat možnost implementování originálních obrázků do souboru, který se vyexportuje. Save images bude tedy zapnutý.

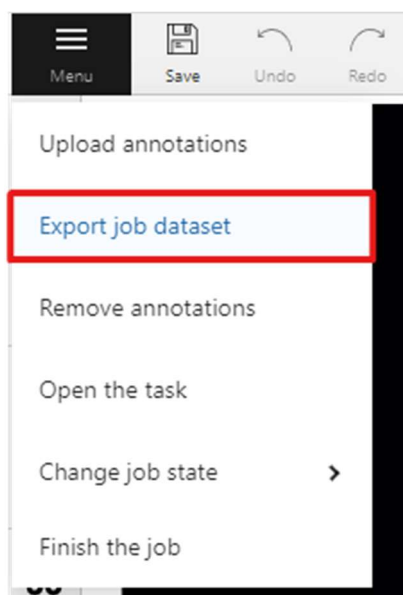
Dále je možnost souboru přiřadit libovolný název. Toto pole se však nemusí vyplňovat a pojmenování vygeneruje CVAT automaticky na základě číslování úkolu a data vytvoření projektu.

Formát souboru je ZIP, tedy komprimovaný archiv. Tímto způsobem dojde ke znatelnému zmenšení velikosti dat. Velkou výhodou tohoto formátu je bezeztrátová komprese pomocí algoritmů, a tak nedojde k možnému znehodnocení anotačních dat. Pro jednoduchost byl zvolen název souboru paznehty.

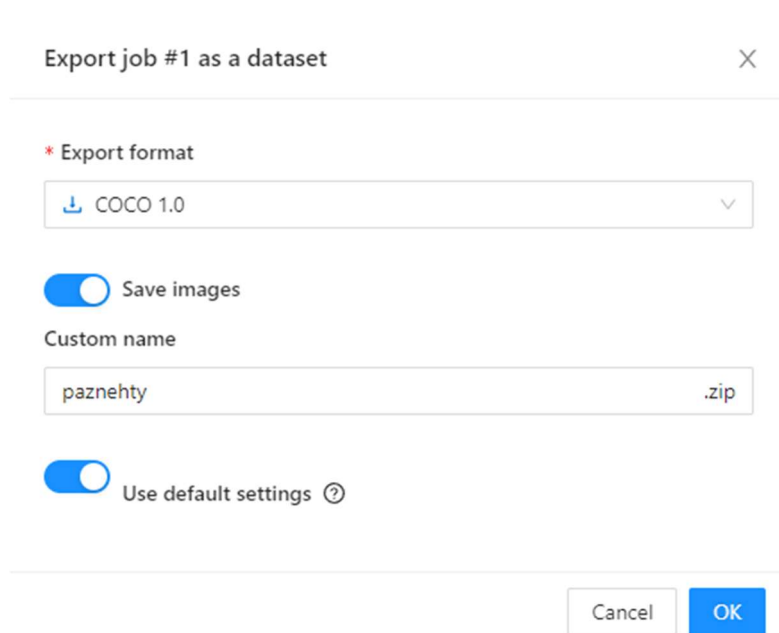
Poslední možností je vybrat výchozí nastavení. Přesněji se jedná o možnost výběru uložště, kam chceme exportovaný dataset nahrát. Na výběr je pouze lokální uložště, tedy disk počítače, na kterém pracujeme, nebo cloudové uložště, které se dále musí specifikovat. V práci je použito lokální uložště.

Nabídku potvrdíme tlačítkem OK. Následně dojde ke stažení souboru ZIP, jelikož software CVAT funguje v prohlížečovém rozhraní.

Tímto je dataset zhotovený a připravený na další využití. Archiv obsahuje dvě složky, z nichž v té s názvem images jsou uloženy všechny snímky a složka annotations obsahuje soubor *instances_default.json* zahrnující veškeré vytvořené anotace.



Obrázek 3.8: Tlačítko pro export datasetu



Obrázek 3.9: Nabídka možností pro exportování

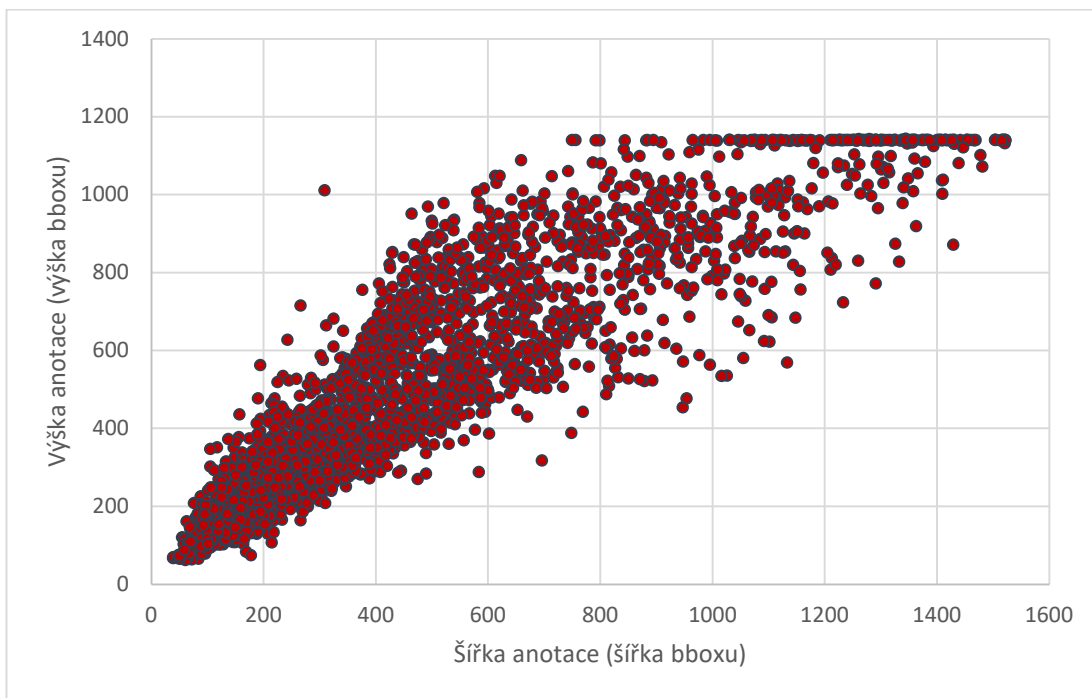
4 Výsledky a diskuse

V rámci bakalářské práce byl vytvořen dataset čítající 1609 snímků, který poslouží jako základ pro trénování a testování CNN. Rozpoznávání paznehtů bylo rozděleno do třech kategorií. Těmito kategoriemi jsou *back* (zadní část paznehtu), *front* (přední část paznehtu) a *side* (boční strana paznehtu). V tabulce 4 je uveden přehled vybraných parametrů anotací ve vytvořeném datasetu. Nejčastěji zastoupenou kategorií je *side* s počtem 1296 anotací. Naopak nejméně se v datasetu vyskytuje kategorie *back* čítající 657 anotací. Velikost anotací je variabilní, jejich šířka se pohybuje v rozmezí 39-1522 pixelů, přičemž průměrná šířka činí 450,94 px, a výška dosahuje hodnot od 62 do 1143 px, průměrná výška anotace je 483,72 px. Tato skutečnost je dána charakterem jednotlivých snímků, kdy některé snímky byly vytvořeny s důrazem na detail v bezprostřední blízkosti termokamery cca 1 m od měřeného paznehtu. V jiných případech se jednalo o snímky, kde vzdálenost od paznehtu se pohybovala cca 6 m, a proto byly výsledné hodnoty pixelů nižší. Tomu odpovídá minimální a maximální plocha anoto- vaných snímků ku celkové ploše snímku, kdy dané hodnoty jsou od 0,10 do 77 % a průměr 10,15 %. Počet bodů u jednotlivých anotací se pohybuje v intervalu od 34 do 852, což je dáno tvarovou složitostí jednotlivých anotací a její plošné velikosti. Ta- bulka 4 představuje přehled vybraných charakteristik popisujících veškeré vytvořené anotace s ohledem na dané kategorie.

Tabulka 4: Přehled vybraných parametrů anotací ve vytvořeném datasetu

	Všechny anotace	boční strana paznehtu	přední část paznehtu	zadní část paznehtu
počet anotací	2946	1296	993	657
min šířka	39	76	63	39
max šířka	1522	1521	1522	1469
prům šířka	450,94	520,69	489,53	255,03
min výška	62	65	83	62
max výška	1143	1141	1141	1143
prům výška	483,72	495,67	577,32	318,67
min plocha	0,10%	0,17%	0,24%	0,10%
max plocha	77,69%	67,57%	77,69%	69,91%
prům plocha	10,15%	10,56%	13,68%	3,98%
min počet bodů	34	40	34	36
max počet bodů	852	636	734	852
prům počet bodů	185,46	192,54	187,26	168,79

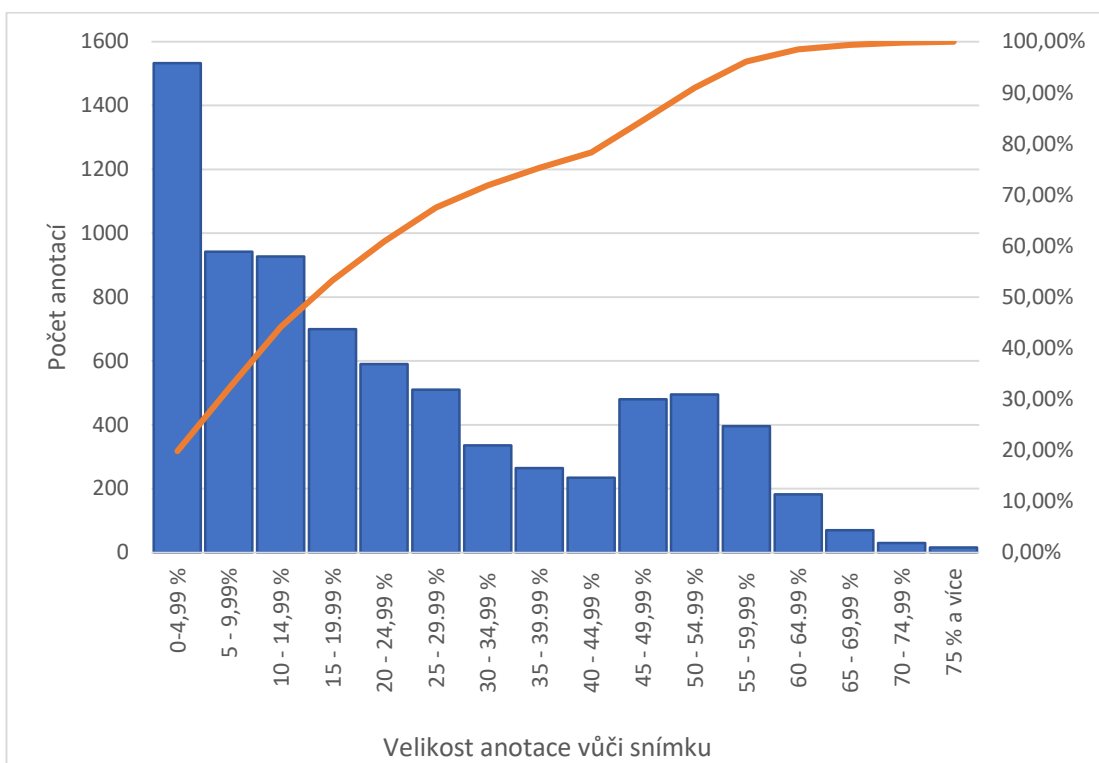
Na grafu 1 je vyobrazeno zastoupení velikosti jednotlivých anotací. Z grafu je patrné, že nejčtenější výskyt mají anotace do velikosti 800x600 px, poměr stran anotací se převážně pohybuje v rozmezí 3:1 až 1:2, kdy první údaj představuje vertikální stranu, tedy výšku, a druhý údaj horizontální stranu, potažmo šířku anotace. Velikost maximální výšky bboxu u anotací je omezena formátem snímku 1200x1600 px.



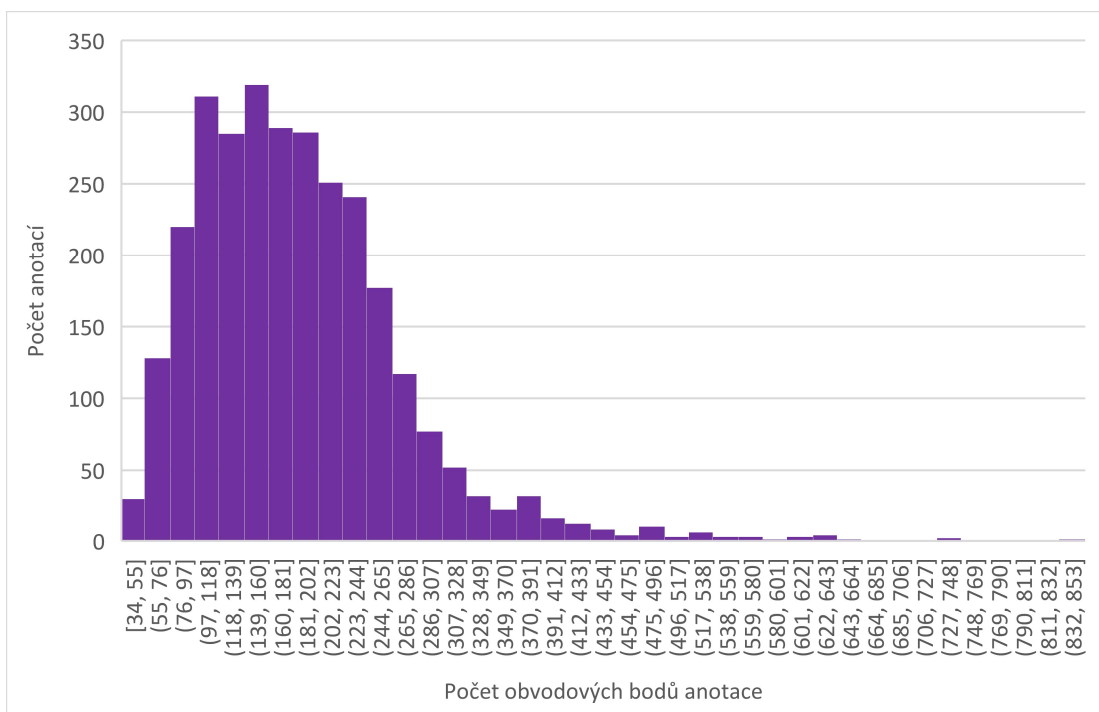
Graf 1: Znázornění velikostí vytvořených anotací

Graf 2 vyobrazuje poměrné zastoupení velikostí anotace ku velikosti snímku. Ze zpracovaných dat vyplývá vysoké zastoupení menších anotací, více než 50 % ze všech anotací mají plochu zabírající méně než 20 % snímku. Velkých anotací o ploše než 45 a více % snímku je v datasetu pouze 15 %.

Z grafu 3 vyplývá, že nejvyšších četností výskytu v datasetu dosahují anotace obsahující 76 až 244 obvodových bodů, 454 bodů a více má již malé množství anotací.



Graf 2: Četnost velikosti anotací vůči snímkům



Graf 3: Četnost anotací dle počtu obvodových bodů

Proces tvorby anotací je časově velmi náročný, jednoduchost ovládání a intuitivní prostředí anotačního softwaru napomáhají k dobrému porozumění a optimálnímu užívání, jelikož si lze snadno osvojit potřebné postupy a není potřeba informace vyhledávat. Při práci s jednotlivými anotačními nástroji vzestupuje otázka, jak dosáhnout jejich spuštění. Některé fungují na webovém rozhraní, jiné je zapotřebí nainstalovat do lokálního úložiště počítače. K funkci softwaru CVAT nebo Label Studio na lokálním úložišti je zapotřebí dalších podpůrných programů jako je například Docker. Jedná se o virtualizaci serverového rozhraní na úložišti počítače, kde se tyto aplikace zpřístupňují pomocí webového prohlížeče přes příslušné adresy s porty. Instalace Dockeru může pro řadu uživatelů představovat překážku z důvodu potřeby zprovoznění virtualizace v operačním systému. CVAT je instalován v několika krocích pomocí příkazového řádku Git Bash, kde je nutností mít alespoň základní zkušenosti s programováním. Další problém nastává při instalaci více než jednoho softwaru. Většina nastavuje defaultní číslo portů stejné, tudíž musí dojít ke změně jednoho z nich, aby rozhraní a tok informací fungovalo správně.

Dalším aspektem je podpora formátů, které software importuje a následně i exportuje. Velké zastoupení uživatelů raději přistoupí k možnosti výběru méně přívětivého prostředí za účelem využití potřebných formátů. Nelze říci, který formát je lepší, jelikož každý je stavěný na trochu odlišné bázi a jejich uplatnění nalezneme ve specifických případech. Jednotlivé struktury lze spíše subjektivně popsat jako přehlednější a smysluplnější.

Nástroje implementované v jednotlivých aplikacích se mohou lišit, nicméně principem je stále ohraničit objekty. Po vytvoření datasetu a následném učení sítě na určité tvary a vlastnosti objektů lze natrénovaný model aplikovat pro tvorbu dalších anotací, kde již není potřeba lidského vstupu.

Závěr

Technologie počítačového vidění a hlubokého učení je aplikována v širokém spektru oborů včetně průmyslu, dopravy, zdravotnictví či zemědělství. V dnešní době se také čím dál více vyskytuje v oblastech každodenního života. Obrazová data jsou standardně získávána pomocí RGB snímačů, avšak některé eventuality jako například záněty končetin u skotu jsou jim skryty. V takových situacích je vhodné použít zařízení umožňující zaznamenávat elektromagnetické záření mimo rozsah viditelného spektra, jako jsou například termokamery. Tato práce je zaměřena na přípravu datasetů termosnímků paznehtů skotu a následné využití v technologiích neuronové sítě.

Při vytváření vhodného datasetu je nutností dbát velkou pozornost přesnosti anotování objektů. Kvalitní výsledek je velice cennou složkou strojového učení a obsáhlost hraje klíčovou roli v pochopení situací. Softwaru pro přípravu trénovacích dat pro CNN je značné množství, tato práce představuje nejoblíbenější z nich a více se věnuje aplikaci softwaru CVAT.

Výsledné anotace mohou být generovány do různých formátů, ze kterých k nejvýznamnějším patří MSCOCO a Pascal VOC, jenž jsou v rámci práce blíže popsány. Klíčovými vlastnostmi těchto formátů je jednoduchost a srozumitelnost, liší se však ve struktuře, přičemž samotná volba formátu se odvíjí od vybrané konvoluční neuronové sítě.

Seznam použité literatury

2012 (VOC2012) Development Kit. [online] host.robots.ox.ac.uk [cit. 19.2.2023]. Dostupné z: <http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>

Aghdam, H., Heravi, J.E. (2017). *Guide to Convolutional Neural Networks*. Springer International Publishing, Tarragona. ISBN 978-3-319-57550-6

Aljabri, M., AlAmir, M., AlGhamdi, M. et al. (2022). *Towards a better understanding of annotation tools for medical imaging: a survey*. *Multimed Tools Appl* 81, 25877–25911. <https://doi.org/10.1007/s11042-022-12100-1>

Andersen, S. (2020). *Everything about CMYK colors*. [online] LimePack Blog [cit. 20.2.2023]. Dostupné z: <https://www.limepack.eu/blog/everything-about-cmyk-colours-2/>

Artuffo, F. (2019). *What is CMYK? What is it used for (in the printing industry)?*. [online] Oppaca [cit. 11.1.2023]. Dostupné z: <https://www.oppaca.com/en/blog/font-color/what-cmyk-what-it-used-printing-industry>

Boesch, G. (2023). *CVAT: Computer Vision Annotation Tool – 2023 Guide*. [online] Viso.ai [cit. 12.1.2023]. Dostupné z: <https://viso.ai/computer-vision/cvat-computer-vision-annotation-tool/>

Boesch, G. (2023). *Image Annotation: Best Software Tools and Solutions in 2023*. [online] Viso.ai [cit. 12.1.2023]. Dostupné z: <https://viso.ai/computer-vision/image-annotation/>

Boesch, G. (2023). *LabelImg for Image Annotation*. [online] Viso.ai [cit. 12.1.2023]. Dostupné z: <https://viso.ai/computer-vision/labelimg-for-image-annotation/>

Brownlee, J. (2019). *a Gentle Introduction to the ImageNet Challenge (ILSVRC)*. [online] Machine Learning Mastery [cit. 22.1.2023]. Dostupné z: <https://machinelearningmastery.com/introduction-to-the-imagenet-large-scale-visual-recognition-challenge-ilsvrc/>

Cocodataset.org (2023). *Data format*. [online] [cit. 19.1.2023]. Dostupné z: <https://cocodataset.org/#format-data>

Datagen.tech (2023). *Labelme: the Basics (Plus a Quick Image Segmentation Tutorial)*. [online] [cit. 25.2.2023]. Dostupné z: <https://datagen.tech/guides/image-annotation/labelme/>

-
- Datagen.tech (2023). *MS COCO Dataset: Using it in Your Computer Vision Projects*. [online] [cit. 13.2.2023]. Dostupné z: <https://datagen.tech/guides/image-datasets/ms-coco-dataset-using-it-in-your-computer-vision-projects/>
- Devopedia (2021). *ImageNet. Version 16, April 7*. Accessed 2023-02-11. [online] Devopedia [cit. 25.2.2023]. Dostupné z: <https://devopedia.org/imagenet>
- Dutta, A. (2018). *VGG Image Annotator*. [online] Visual Geometry Group blog [cit. 28.2.2023]. Dostupné z: <https://www.robots.ox.ac.uk/~vgg/blog/vgg-image-annotator.html>
- Dutta, A. a Zisserman, A. (2019). *The VIA Annotation Software for Images, Audio and Video*. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19). Association for Computing Machinery, New York, NY, USA, 2276–2279. <https://doi.org/10.1145/3343031.3350535>
- Everingham, M., Winn, J. (2012). *The PASCAL Visual Object Classes Challenge*
- Exxact (2022). *Intro and Short Tutorial on Data Annotation with Label Studios*. [online] [cit. 25.3.2023]. Dostupné z: <https://www.exxactcorp.com/blog/Deep-Learning/Intro-and-Short-Tutorial-on-Data-Annotation-with-Label-Studios>
- Fák, T. (2014). *Onemocnění paznehtů a jejich vliv na užitkovost a reprodukci krav*. Diplomová práce, Jihočeská univerzita v Českých Budějovicích, Zemědělská fakulta.
- Forsyth, D.A. a Ponce, J. (2011). *Computer Vision a Modern Approach*. Druhé vydání. Prentice Hall, pp.792. ISBN 978-0136085928.
- Gagny, E. (2023). *Label Studio for Anomaly Detection at Scale*. [online] Medium [cit. 30.3.2023]. Dostupné z: <https://medium.com/total-digital-factory/label-studio-for-anomaly-detection-at-scale-7b50dc4e46e9>
- Goodfellow, I., Bengio, Y., Courville, A. (2016). *Deep learning*. MIT Press, ISBN 978-0-262-03561-3
- Guha, R. (2021). *RGB v. HSV for Computer Vision*. [online] rehanguha github io [cit. 21.4.2023]. Dostupné z: <https://rehanguha.github.io/articles/rbg-vs-hsv-for-computer-vision>
- Hassaballah, M. a Awad, A.I. (2020). *Deep learning in computer vision principles and applications*. CRC Press, Taylor & Francis Group, Boca Raton. ISBN 978-1-138-54442-0.
- Howlett, CH. T. (2017). *Why i Love HSV, and Why It's Totally Useless*. [online] Howlett studios [cit. 26.2.2023]. Dostupné z: <https://howlettstudios.com/articles/2017/5/6/the-problem-with-hsv>
-

-
- Humansintheloop.org, (2020). *Top 4 Annotation Tools for Computer Vision*. [online] [cit. 21.1.2023]. Dostupné z: <https://humansintheloop.org/top-4-annotation-tools-for-computer-vision/>
- Hunter, J., Dále, D., Firing, E., Droettboom, M. (2012). *Choosing Colormaps in Matplotlib*. [online] Matplotlib [cit. 21.3.2023] Dostupné z: <https://matplotlib.org/stable/tutorials/colors/colormaps.html>
- Immersivelimit.com (2019). *Create COCO Annotations From Scratch*. [online] [cit. 18.2.2023]. Dostupné z: <https://www.immersivelimit.com/tutorials/create-coco-annotations-from-scratch>
- Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B. (2022). *a Review of Yolo Algorithm Developments*, Procedia Computer Science, Volume 199, Pages 1066-1073, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2022.01.135>
- Kang, H.-C., Han, H.-N., Bae, H.-C., Kim, M.-G., Son, J.-Y., & Kim, Y.-K. (2021). *HSV Color-Space-Based Automated Object Localization for Robot Grasping without Prior Knowledge*. [online] Applied Sciences, 11(16), 7593. [cit. 17.1.2023]. Dostupné z: <https://doi.org/10.3390/app11167593>
- Karatas, G. (2021). *Data Annotation in 2023: Why it matters & Top 8 Best Practices*. [online] Research.aimultiple [cit. 9.2.2023]. Dostupné z: <https://research.aimultiple.com/data-annotation/>
- Khandelwal, R. (2019). *COCO and Pascal VOC data format for Object detection*. [online] Towardsdatascience [cit. 26.2.2023]. Dostupné z: <https://towardsdatascience.com/coco-data-format-for-object-detection-a4c5eaf518c5>
- Knížková, I. a Kunc, P. (2020). *Infračervená termografie a její využití*
- Knížková, I. a Kunc, P. (2021). *Infračervená termografie. Specifika a správné postupy měření živých objektů*. Výzkumný ústav živočišné výroby, v. v. i., Praha Uhřetěves. ISBN 978-80-7403-258-5
- Kolář, M. (2020). *Computer vision with active learning*. Disertační práce, Vysoké učení technické v Brně, Fakulta informačních technologií.
- Kovalčík, V. (2020). *sRGB, ProPhoto RGB a další – vyznáte se v barevných prostorech?*. [online] Milujeme fotografii [cit. 18.1.2023]. Dostupné z: <https://www.milujemefotografii.cz/srgb-prophoto-rgb-a-dalsi-vyznate-se-v-barevných-prostorech>
- Kundu, R. (2023). *YOLO: Algorithm for Object Detection Explained [+Examples]*. [online] V7labs [cit. 12.2..2023]. Dostupné z: <https://www.v7labs.com/blog/yolo-object-detection>
-

Liu, H., Li, W., Li, Z., Wang, Y., Gui, G. (2019). *Deep Learning Based Adversarial Images Detection*. In: Gui, G., Yun, L. (eds.): *Advanced Hybrid Information Processing*. ADHIP 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 301. Springer, Cham.

Liubimov, N. (2020). *Introducing Label Studio, a swiss army knife of data labeling*. [online] Towards Data Science [cit. 27.3.2023]. Dostupné z: <https://towardsdatascience.com/introducing-label-studio-a-swiss-army-knife-of-data-labeling-140c1be92881>

Lunden, I. (2022). *V7 snaps up \$33M to automate training data for computer vision AI models*. [online] Techcrunch [cit. 16.2.2023]. Dostupné z: <https://techcrunch.com/2022/11/28/v7-labs-computer-vision-ai/>

Mangla, P. (2022). *Computer Vision and Deep Learning for Agriculture*. [online] Pyimagesearch [cit. 16.2.2023]. Dostupné z: <https://pyimagesearch.com/2022/08/10/computer-vision-and-deep-learning-for-agriculture/>

Nelson, J. (2020) *LabelImg for Labeling Object Detection Data*. [online] Blog.roboflow [cit. 23.2.2023]. Dostupné z: <https://blog.roboflow.com/labelimg/>

Nelson, J. (2021). *Your Comprehensive Guide to the YOLO Family of Models*. [online] Blog.roboflow [cit. 13.2.2023]. Dostupné z: <https://blog.roboflow.com/guide-to-yolo-models/>

Patel, I. (2022). *Computer Vision and Deep learning for Agriculture*. [online] LinkedIn [cit. 26.12.2022]. Dostupné z: https://www.linkedin.com/pulse/computer-vision-deep-learning-agriculture-ashwani-patel?trk=pulse-article_more-articles_related-content-card

Petrosyan, T. (2021). *Introduction to the COCO Dataset*. [online] Opencv [cit. 17.2.2023]. Dostupné z: <https://opencv.org/introduction-to-the-coco-dataset/>

Pokhrel, S. (2020). *Image Data Labelling and Annotation — Everything you need to know*. [online] Towardsdatascience [cit. 7.2.2023]. Dostupné z: <https://towardsdatascience.com/image-data-labelling-and-annotation-everything-you-need-to-know-86ede6c684b1>

Pro zjišťování pohody a zdraví hospodářských zvířat. Výzkumný ústav živočišné výroby, v.v.i., Zemědělský svaz ČR – Česká technologická platforma pro zemědělství, Praha. ISBN 978-80-7403-243-1

Redmon, J. C., Farhadi, A. (2023). *YOLO: Real-Time Object Detection*. [online] Pjred-die [cit. 6.3.2023]. Dostupné z: <https://pjred-die.com/darknet/yolo/>

-
- Rosebrock, A. (2017). *ImageNet: VGGNet, ResNet, Inception, and Xception with Keras*. [online] Pyimagesearch [cit. 9.2.2023]. Dostupné z: <https://pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception-xception-keras/>
- Rosebrock, A. (2021). *OpenCV Color Spaces*. [online] Pyimagesearch [cit. 25.2.2023]. Dostupné z: <https://pyimagesearch.com/2021/04/28/opencv-color-spaces-cv2-cvcolor/>
- Russ, J.C. (2002). *The Image Processing Handbook*. CRC Press, Boca Raton. ISBN 0-8493-1142-X
- Russell, B.C., Torralba, A., Murphy, K.P. et al. (2008). *LabelMe: a Database and Web-Based Tool for Image Annotation*. *Int J Comput Vis* 77, 157–173. <https://doi.org/10.1007/s11263-007-0090-8>
- Sciendo.com (2022). *XML files for an increased visibility*. [online] [cit. 14.1.2023]. Dostupné z: <https://sciendo.com/blog/xml-files-for-an-increased-visibility>
- Sensing.konicaminolta.us, (2023). *Identifying Color Differences Using L*a*b* or L*C*H* Coordinates*. [online] [cit. 18.2.2023]. Dostupné z: <https://sensing.konicaminolta.us/us/blog/identifying-color-differences-using-l-a-b-or-l-c-h-coordinates/>
- Shah, D. (2023). *COCO Dataset: All You Need to Know to Get Started*. [online] V7labs [cit. 15.2.2023]. Dostupné z: <https://www.v7labs.com/blog/coco-dataset-guide>
- Sharma, A. (2022). *Understanding a Real-Time Object Detection Network: You Only Look Once (YOLOv1)*. [online] Pyimagesearch [cit. 19.1.2023]. Dostupné z: <https://pyimagesearch.com/2022/04/11/understanding-a-real-time-object-detection-network-you-only-look-once-yolov1/>
- Solawetz, J. (2022). *What is CVAT (Computer Vision Annotation Tool)?*. [online] Blog.roboflow [cit. 13.1.2023]. Dostupné z: <https://blog.roboflow.com/cvat/>
- Solawetz, J., Francoco (2023). *What is YOLOv8? The Ultimate Guide*. [online] Blog.roboflow [cit. 28.2.2023]. Dostupné z: <https://blog.roboflow.com/whats-new-in-yolov8/>
- Syndorenko, I. (2021). *CVAT Annotation Tool: Review, Definition, Usage*. [online] Labelyourdata [cit. 19.1.2023]. Dostupné z: <https://labelyourdata.com/articles/what-is-cvat-annotation-tool>
- Šonka, M., Hlaváč, V., Boyle, R. (2013). *Image processing, Analysis, and Machine Vision*. THOMSON. ISBN 978-9386858146
- Thermascan.co.uk, (2019). *Understanding Thermal Palettes*. [online] [25.2.2023]. Dostupné z: <https://www.thermascan.co.uk/blog/thermal-palettes>
-

-
- Torralba, A., Russell, B.C. a Yuen J. (2010). *LabelMe: Online Image Annotation and Applications*. Proceedings of the IEEE, vol. 98, no. 8, pp. 1467-1484, Aug. 2010, <http://dx.doi.org/10.1109/JPROC.2010.2050290>
- Tyson, M. (2022). *What is JSON? The universal data format*. [online] InfoWorld [cit. 27.1.2023]. Dostupné z: <https://www.infoworld.com/article/3222851/what-is-json-a-better-format-for-data-exchange.html>
- Vseoprumyslu.cz, (2020). *Robotika a umělá inteligence pro pěstování*. [online] [cit. 16.1.2023]. Dostupné z: <https://www.vseoprumyslu.cz/digitalizace/umela-inteligence/robotika-a-umela-inteligence-pro-pestovani.html>
- Wordnet.princeton.edu (2023). *What is WordNet?*. [online] [cit. 13.1.2023]. Dostupné z: <https://wordnet.princeton.edu/>
- Yao, R. (2020). *Pixel Perfect: V7 Labs Automates Image Annotation for Deep Learning Models*. [online] Blogs.nvidia [cit. 5.2.2023]. Dostupné z: <https://blogs.nvidia.com/blog/2020/08/28/v7-labs-image-annotation/>
- Zelazko, A. (2022). *RGB colour model*. [online] Encyclopedia Britannica [cit. 26.2.2023]. Dostupné z: <https://www.britannica.com/science/RGB-colour-model>
-

Seznam obrázků

Obrázek 1.1: Struktura CNN.....	12
Obrázek 1.2: Pracovní prostředí v softwaru CVAT.....	18
Obrázek 1.3: Pracovní prostředí softwaru V7.....	19
Obrázek 1.4: Pracovní prostředí softwaru VIA.....	21
Obrázek 1.5: Pracovní prostředí softwaru LabelMe	22
Obrázek 1.6: Pracovní prostředí softwaru LabelImg	23
Obrázek 1.7: Pracovní prostředí softwaru Label Studio	24
Obrázek 1.8: Základní datová struktura MS COCO	26
Obrázek 1.9: Struktura anotace MS COCO	27
Obrázek 1.10: Struktura anotace s větším počtem bodů MS COCO	27
Obrázek 1.11: Struktura Pascal VOC	29
Obrázek 1.12: Princip hierarchie ImageNetu.....	32
Obrázek 1.13: RGB prostor se spojením barev.....	34
Obrázek 1.14: CMYK prostor se skládáním.....	35
Obrázek 1.15: HSV prostor se směrem hodnot.....	36
Obrázek 1.16: L*a*b* prostor.....	37
Obrázek 1.17: Ironbow termosnímek s inferno škálou	39
Obrázek 1.18: Rainbow snímek s jet škálou	40
Obrázek 1.19: White hot termosnímek se škálou	40
Obrázek 3.1: Klasický snímek paznehtu.....	42
Obrázek 3.2: Nabídka vytvoření nového projektu	43
Obrázek 3.3: Nabídka vytvoření nového úkolu	44
Obrázek 3.4: Pracovní plocha v softwaru CVAT	45
Obrázek 3.5: Nabídky pravého panelu.....	48
Obrázek 3.6: Postup obkreslování vlevo, tlačítko pro potvrzení výběru vpravo	50
Obrázek 3.7: Ukončený výběr	50
Obrázek 3.8: Tlačítko pro export datasetu	52
Obrázek 3.9: Nabídka možností pro exportování	52

Seznam tabulek

Tabulka 1: Nástroje softwaru CVAT	46
Tabulka 2: Nástroje softwaru CVAT	47
Tabulka 3: Nástroje softwaru CVAT	47
Tabulka 4: Přehled vybraných parametrů anotací ve vytvořeném datasetu.....	53

Seznam grafů

Graf 1: Znázornění velikostí vytvořených anotací.....	54
Graf 2: Četnost velikosti anotací vůči snímkům.....	55
Graf 3: Četnost anotací dle počtu obvodových bodů.....	55

Seznam použitých zkratek

CNN	Convolutional Neural Networks
CVAT	Computer Vision Annotation Tool
VIA	VGG Image Annotator
XML	Extensible Markup Language
W3C	World Wide Web Consortium
SGML	Standard Generalized Markup Language
HTML	Hypertext Markup Language
JSON	JavaScript Object Notation
MS COCO	Microsoft Common Objects in Context
VOC	Visual Object Classes
YOLO	You Only Look Once
ResNet	Residual neural network
CSPNet	Cross Stage Partial Network
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
RGB	red green blue
CMYK	cyan magenta yellow key
HSV	hue saturation value
LAB	L^* ; a^* ; b^*
HC	High Contrast
OpenCV	Open Source Computer Vision Library
URL	Uniform Resource Locator
