

Jihočeská univerzita v Českých Budějovicích
Ekonomická fakulta
Katedra aplikované matematiky a informatiky

Teze diplomové práce

Segmentace snímků z termokamery aplikací vybraných konvolučních neuronových sítí

Vypracoval: Bc. Lukáš Beneda

Vedoucí práce: doc. RNDr. Petr Bartoš, Ph.D.

České Budějovice

2023

Klíčová slova:

Termografie, neuronové sítě, pazneht skotu, YOLOv7, Mask R-CNN, Mask DINO, segmentace instancí

Anotace:

Tato práce se zabývá problémem segmentace instancí paznehtu skotu na termografických snímcích pomocí neuronových sítí. Cílem této práce bylo otestovat několik vybraných řešení a vyhodnocení jejich výsledků. Základem práce je rešerše existujících řešení, příprava datasetu, výběr modelů neuronových sítí a ohodnocení výsledků jednotlivých modelů. Práce popisuje postup prací a závěrem jsou porovnány jednotlivé výsledky a vyhodnoceno nejlepší řešení. Výsledkem práce je vytvořený dataset termografických snímků paznehtů a 3 otestované modely neuronových sítí, ze kterých 2 jsou velmi dobře použitelné pro řešení daného problému.

Úvod

Strojové učení je jednou z oblastí umělé inteligence, která přináší možnost automatického osvojení konceptů a znalostí bez nutnosti explicitního naprogramování těchto poznatků. Klasické strojové učení je v posledních letech překonáváno technologií deep learningu. Deep learning využívá technologie neuronových sítí a pokročilých učících algoritmů. Proces přípravy této technologie pro konkrétní úlohu se skládá ze dvou částí, kdy v první z nich probíhá samotné učení na předpřipravených datech a v druhé ověření efektivity a přesnosti pomocí nových, pro neuronovou síť neznámých dat. (Dargan, Kumar, Ayyagari, & Kumar, 2020) (Alzubaidi et al., 2021) (Dong, Wang, & Abbas, 2021)

V dnešní době je v komunitě zabývající se neuronovými sítěmi řešena dlouhá řada výzev od detekce a rozpoznávání dopravních značek (Zhang, Wang, Song, Yan, & Qi, 2022), přes rozlišování jednotlivých kamenů v sypkých materiálech (Yang et al., 2021), rozpoznávání obličejů (Sanchez-Moreno et al., 2021), různé detekce poškození (Pham, Pham, & Dang, 2020) (Wen, Huang, & Guo, 2021), počítání objektů na leteckých snímcích (Wu et al., 2021) (Osco et al., 2021), až po různé výzvy v medicíně (Li, Du, Luo, Wang, & Ding, 2023) (Brodzicki, Jaworek-Korjakowska, Kleczek, Garland, & Bogyo, 2020) (Long, Peng, Song, Xia, & Sang, 2021) (A. Al-Falluji, Dalaf Katheeth, & Alathari, 2021) (Zaridis et al., 2023) nebo právě v zemědělství (Santos, de Souza, dos Santos, & Avila, 2020) (Sarwar, Griffin, Rehman, & Pasang, 2021) (Maksimovich, Fedorov, & Kalichkin, 2023).

Tato diplomová práce se zabývá využitím technologie neuronových sítí pro detekci paznehtu skotu, zejména u krav, v termografických snímcích pro následné automatické vyhodnocení anomálií a případné detekce onemocnění paznehtu, které mohou mít negativní vliv na zdraví zvířete nebo na délku jeho života.

Cíl této práce je připravit a porovnat několik modelů neuronových sítí určených k segmentaci instancí. První část se zabývá samotnou problematikou termografických snímků a aplikace této technologie v praxi. V další kapitole je shrnuta problematika neuronových sítí, a především konvoluční neuronové sítě a jejich aplikace. Ve čtvrté kapitole jsou zmapovány možnosti řešení pro úlohu segmentace obrazu, včetně segmentace obrazu pomocí neuronových sítí. Druhá polovina práce se zabývá přípravou datasetu pro učení sítí na rozpoznávání kravského paznehtu. Následně je popsáno učení vybraných modelů neuronových sítí, a nakonec jsou jednotlivé vybrané modely porovnány z hlediska úspěšnosti, kvality nebo rychlosti inferencí.

Cíl práce

Práce se zabývá problematikou zpracování termálních snímků pomocí vybraných konvolučních neuronových sítí. Cílem teoretické části je podrobná rešerše oblasti využití vyspělých technik počítačového vidění v bezkontaktní termografii. Dílčím cílem práce je vytvoření kvalitního datasetu anotací k termografickým snímkům pořízených v prostředí stájového chovu skotu, zaměřených na oblast dolních končetin, který bude využit pro potřeby trénování neuronových sítí. Hlavním cílem práce je experimentálně otestovat vybrané pokročilé metody hlubokého učení pro segmentaci obrazu a jejich výsledky porovnat a vybrat nejlepší řešení.

Metodika

K vytvoření anotací k pořízeným snímkům jsme použili nástroj CVAT – Computer vision annotation tool, který poskytuje různé anotační nástroje včetně polygonální masky, kterou jsme pro tento dataset využili. Vytvořený dataset obsahuje 1227 snímků s 2479 označenými objekty.

Z provedené rešerše vyšli jako nejvhodnější k trénování na vytvořeném datasetu sítě YOLOv7 (Wang, Bochkovski, & Liao, 2022), Mask R-CNN (He, Gkioxari, Dollar, & Girshick, 2020) a Mask DINO (Li et al., 2022), které jsme experimentálně otestovali na pokusné sadě snímků. Otestování zahrnovalo přeučení modelů na trénovací části našeho datasetu a následné ověření efektivnosti a přesnosti na validační části datasetu.

Při vyhodnocování výsledků jsou sledovány řady metrik, které se uplývají dané úloze. Využívají se také při porovnávání výkonu různých architektur mezi sebou na obecně známých datasetech.

Mezi základní metriky se řadí Přesnost (Accuracy), jenž vyjadřuje poměr správně označených vzorků ku celkovému počtu vyhodnocovaných vzorků. (Alzubaidi et al., 2021) Citlivost (Recall nebo Sensitivity) je vyjadřována pro každou třídu zvlášť, jde o poměr mezi správně označenými vzorky patřící do dané skupiny ku celkovému počtu vzorků, které byly sítě označeny, že do skupiny patří. Vyjadřuje správnost pozitivních označení. (Alzubaidi et al., 2021) Účinnost nebo specifická (Specificity) je počítána pro každou třídu zvlášť a vyjadřuje správnost negativní označení obdobně jako citlivost pro pozitivní označení. (Alzubaidi et al., 2021) Přesnost (Precision) vyjadřuje správnost sítě pro danou třídu. Definována je jako poměr správně označených ku počtu všech, které k dané třídě náleží. (Alzubaidi et al., 2021) Při klasifikaci do více tříd se nejčastěji používá matice záměn, ve které jsou zaneseny správné detekce i záměny mezi třídami. Ideální matice záměn má nenulová čísla pouze na hlavní diagonále.

Některé úlohy jsou ovšem složitější a vyžadují komplexnější metriky. IoU (Intersection over Union) je metrika, která umožňuje porovnání masek nebo bounding boxů. Je vyjádřena jako poměr mezi průnikem a sjednocením dvou oblastí, skutečným označením, které odpovídá realitě, a označením, které vygenerovala síť. (Wu, Sahoo, & Hoi, 2020) Metrikou pro detekce více objektů na jednom snímku je mAP (mean average precision), která kombinuje různé metriky mezi třídami a objekty pomocí průměru. (Wu, Sahoo, & Hoi, 2020) F1 score a J score jsou metriky, které vznikají kombinací základních metrik zmíněných výše. (Alzubaidi et al., 2021)

Výsledky

Na následujících obrázcích 1, 2 a 3 je ukázka vzniklého datasetu. V tabulce 1 je uvedeno shrnutí výsledků z hlediska mAP, která má největší vypovídající hodnotu o výkonu sítě. Síť Mask R-CNN a YOLOv7 se svými výsledky blíží 70 respektive 90 procentům úspěšnosti při predikcích masek. Síť Mask DINO se ukázala jako neideální volba pro zpracování termosnímků. I z hlediska rychlosti inferencí se ukázala síť YOLOv7 jako nejlepší. Na obrázku 4 je vynesena závislost přesnosti predikované masky na míře jistoty predikované třídy sítě YOLOv7, která i při nižší míře jistoty dosahuje vysoké přesnosti. Na obrázku 5 je v grafech uveden vývoj důležitých metrik pro trénování sítě YOLOv7, především vývoj hodnot ztrátových funkcí. V těchto grafech je na vodorovné ose vyneseno číslo iterace a na ose y je uvedena hodnota dané metriky. Z těchto grafů můžeme vyčíst, že nejlepší výsledky podávala síť kolem 150 iterace, ze které jsou parametry modelu uloženy pro následné použití. Jako nejlepší z testovaných bychom tedy vybrali síť YOLOv7.

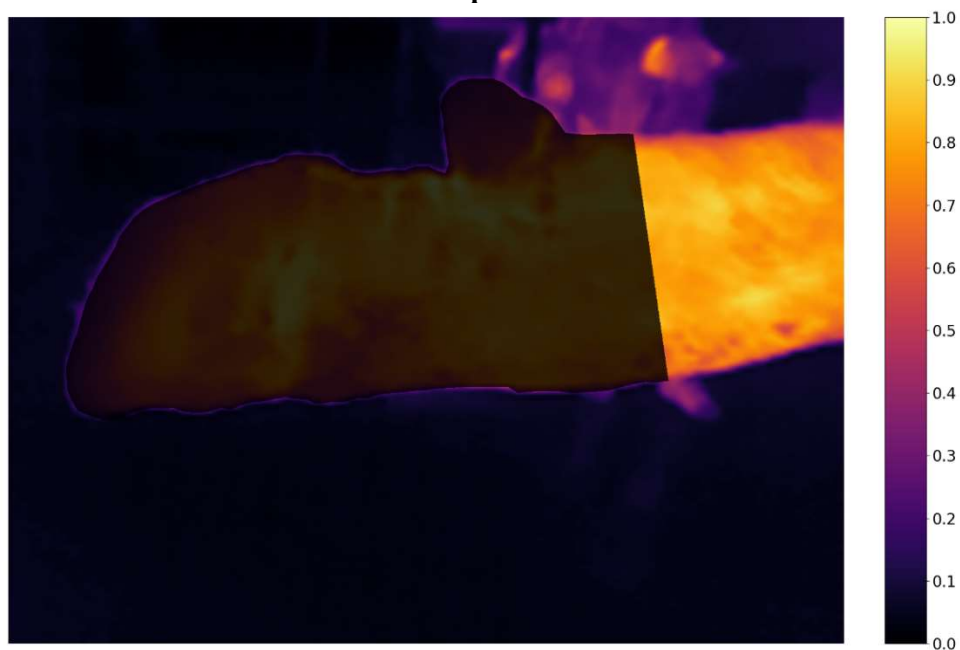
Obrázek 1: Ukázka datasetu – snímek z větší dálky



Obrázek 2: Ukázka datasetu – snímek paznehtu zředu



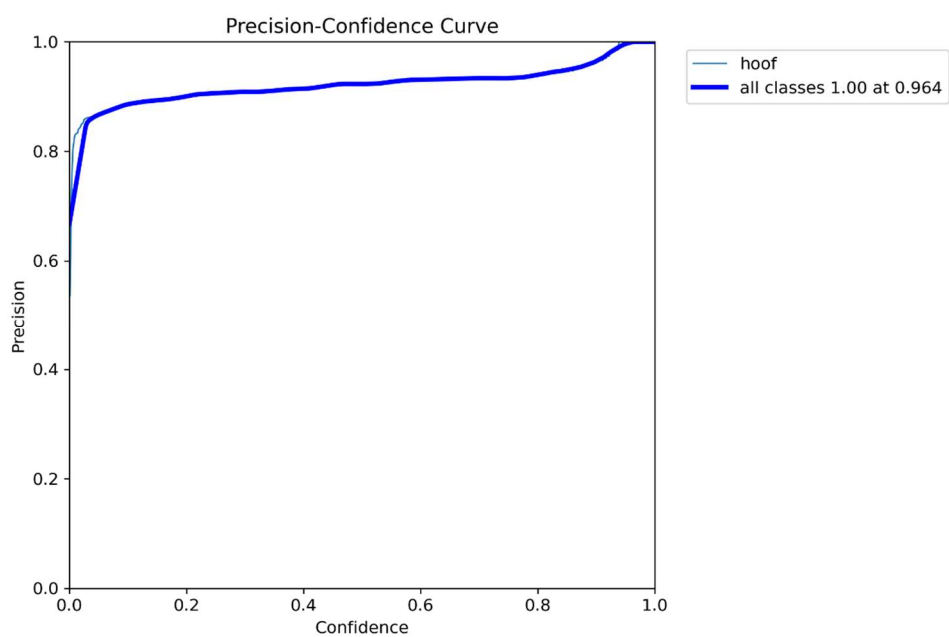
Obrázek 3: Ukázka datasetu – snímek paznehtu z boku



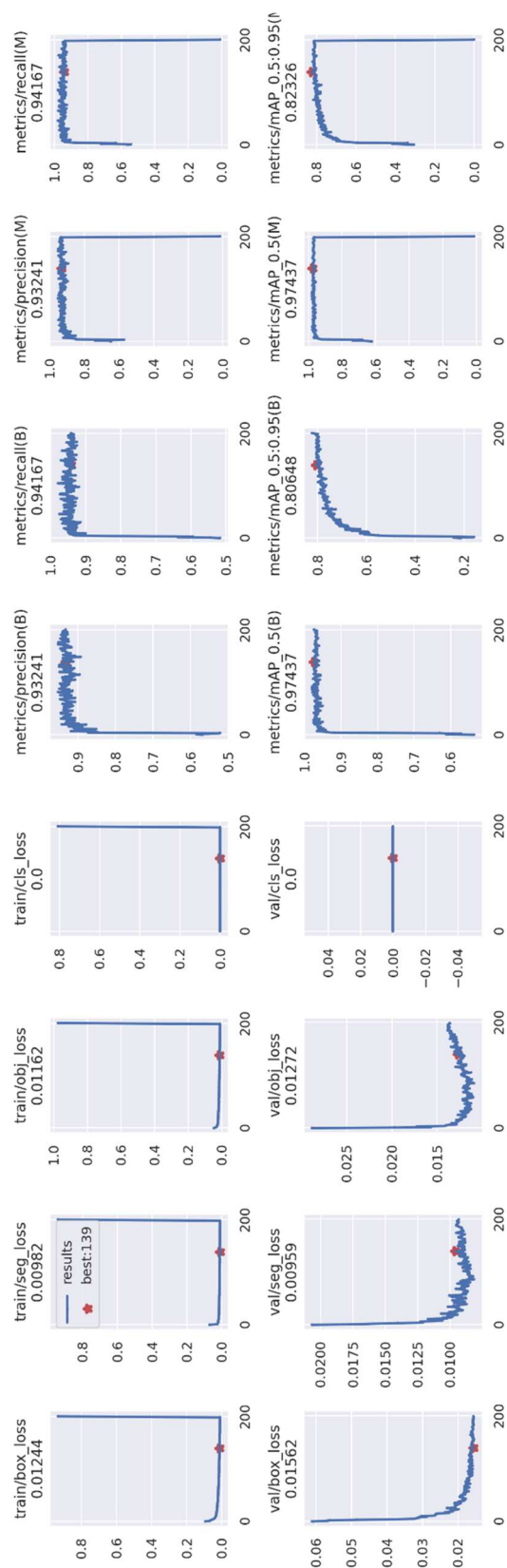
Tabulka 1: Porovnání výsledků

	Box mAP	Mask mAP
Mask R-CNN	0.718	0.715
YOLOv7	0.886	0.889
Mask DINO	0.010	0.361

Obrázek 4: YOLOv7 Mask P-C graf



Obrázek 5: YOLOv7 průběh trénování



Závěr

V této práci jsem se zaměřil na termografické snímky a jejich zpracování pomocí konvolučních neuronových sítí, zejména na úlohu segmentace instancí, která je považována za jeden z náročnějších úkolů.

Cílem práce bylo připravit a otestovat několik modelů konvolučních neuronových sítí, které budou schopny rozeznat na termografickém snímku jednotlivé instance paznehtu skotu. Takto označené instance by mohly být dále využity k další analýze, která by ulehčila diagnostiku zdravotního stavu zvířete.

V první části práce vznikla obecná shrnující rešerše, ze které jsem zjistil, že využití termografických snímků je v medicíně a jiných oborech poměrně běžné. Dále jsem také zjistil, že jejich zpracování pomocí neuronových sítí se nijak zvlášť neliší od klasického zpracování obrazu. Nakonec byly vybrány 3 vhodné modely konvolučních neuronových sítí k natrénování. Mezi vybranými byly Mask R-CNN, YOLOv7 a Mask DINO. V rešerši jsem se seznámil také s několika zajímavými sítěmi, jako je například EVA, které mají ale značně vyšší nároky na hardware.

Ve druhé části práce vznikl kvalitní dataset termografických snímků paznehtů skotu, který obsahuje přes 1200 snímků a anotací ve formátu COCO. Dataset obsahuje snímky z různé vzdálenosti a z různých úhlů. Každý ze snímků má alespoň jednu anotaci. Snímky z větší dálky mají většinou anotací více.

V další části jsem vybrané modely učil na předpřipravených datech a vyhodnocoval jejich spolehlivost. Mezi třemi vybranými modely byly natrénovány síť Mask R-CNN, která dosahovala na necelých 72 % ve standardní metrice mAP pro hodnocení přesnosti masek, YOLOv7, která se ve vyhodnocení pyšní mAP 89 %, a Mask DINO, která bohužel absolutně nesplnila předpoklady a dosáhla na pouhých 36 % mAP. Mohlo zde sehrát svoji roli také to, že Mask DINO při učení využívá velmi velké množství paměti, a proto se při učení využívalo nižší paralelizace zpracování.

Z hlediska rychlosti predikce se osvědčila velmi dobře síť YOLOv7, které stačilo k predikci na termografickém snímku 3,7 ms, což odpovídá přibližně 270 snímkům za vteřinu na naší testovací sadě. S výkonnějším hardwarem by mělo být možné dosáhnout i vyšší frekvence.

Z výzkumu vyšla jako nejlepší volba pro segmentaci instancí paznehtu skotu síť YOLOv7. Špatnou volbou by nebyla ani Síť Mask R-CNN.

Seznam literatur použité v tezích

Plný seznam použité literatury je uveden v textu diplomové práce.

- A. Al-Falluji, R., Dalaf Katheeth, Z., & Alathari, B. (2021). Automatic Detection of COVID-19 Using Chest X-Ray Images and Modified ResNet18-Based Convolution Neural Networks [Online]. 66(2), 1301-1313. <https://doi.org/10.32604/cmc.2020.013232>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., et al. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions [Online]. Journal Of Big Data, 8(1). <https://doi.org/10.1186/s40537-021-00444-8>
- Brodzicki, A., Jaworek-Korjakowska, J., Kleczek, P., Garland, M., & Bogyo, M. (2020). Pre-Trained Deep Convolutional Neural Network for Clostridioides Difficile Bacteria Cytotoxicity Classification Based on Fluorescence Images [Online]. Sensors, 20(23). <https://doi.org/10.3390/s20236713>
- Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2020). A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning [Online]. Archives Of Computational Methods In Engineering, 27(4), 1071-1092. <https://doi.org/10.1007/s11831-019-09344-w>
- Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications [Online]. Computer Science Review, 40. <https://doi.org/10.1016/j.cosrev.2021.100379>
- He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2020). Mask R-CNN [Online]. Ieee Transactions On Pattern Analysis And Machine Intelligence, 42(2), 386-397. <https://doi.org/10.1109/TPAMI.2018.2844175>
- Li, C., Du, R., Luo, Q., Wang, R., & Ding, X. (2023). A Novel Model of Thyroid Nodule Segmentation for Ultrasound Images [Online]. 49(2), 489-496. <https://doi.org/10.1016/j.ultrasmedbio.2022.09.017>
- Li, F., Zhang, H., xu, H., Liu, S., Zhang, L., Ni, L. M., & Shum, H. -Y. (2022). Mask DINO: Towards A Unified Transformer-based Framework for Object Detection and Segmentation [Online]. <https://doi.org/10.48550/arXiv.2206.02777>
- Long, F., Peng, J. -J., Song, W., Xia, X., & Sang, J. (2021). BloodCaps: A capsule network based model for the multiclassification of human peripheral blood cells [Online]. Computer Methods And Programs In Biomedicine, 202. <https://doi.org/10.1016/j.cmpb.2021.105972>
- Maksimovich, K., Fedorov, D., & Kalichkin, V. (2023). Feed Forward Neural Network Modelling for Spring Wheat Crop Forecast [Online]. Agriculture Digitalization And Organic Production, 303-312. https://doi.org/10.1007/978-981-19-7780-0_27
- Oscio, L. P., Marcato Junior, J., Marques Ramos, A. P., de Castro Jorge, L. A., Fatholahi, S. N., de Andrade Silva, J., et al. (2021). A review on deep learning in UAV remote sensing [Online]. International Journal Of Applied Earth Observation And Geoinformation, 102. <https://doi.org/10.1016/j.jag.2021.102456>
- Pham, V., Pham, C., & Dang, T. (2020). Road Damage Detection and Classification with Detectron2 and Faster R-CNN [Online]. 2020 Ieee International Conference On Big Data (Big Data), 5592-5601. <https://doi.org/10.1109/BigData50022.2020.9378027>
- Sanchez-Moreno, A. S., Olivares-Mercado, J., Hernandez-Suarez, A., Toscano-Medina, K., Sanchez-Perez, G., & Benitez-Garcia, G. (2021). Efficient Face Recognition System for Operating in Unconstrained Environments [Online]. Journal Of Imaging, 7(9). <https://doi.org/10.3390/jimaging7090161>
- Santos, T. T., de Souza, L. L., dos Santos, A. A., & Avila, S. (2020). Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association [Online]. Computers And Electronics In Agriculture, 170. <https://doi.org/10.1016/j.compag.2020.105247>
- Sarwar, F., Griffin, A., Rehman, S. U., & Pasang, T. (2021). Detecting sheep in UAV images [Online]. Computers And Electronics In Agriculture, 187. <https://doi.org/10.1016/j.compag.2021.106219>
- Wang, C. -Y., Bochkovskiy, A., & Liao, H. -Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [Online]. <https://doi.org/10.48550/arXiv.2207.02696>
- Wen, H., Huang, C., & Guo, S. (2021). The Application of Convolutional Neural Networks (CNNs) to Recognize Defects in 3D-Printed Parts [Online]. Materials, 14(10). <https://doi.org/10.3390/ma14102575>
- Wu, Q., Feng, D., Cao, C., Zeng, X., Feng, Z., Wu, J., & Huang, Z. (2021). Improved Mask R-CNN for Aircraft Detection in Remote Sensing Images [Online]. Sensors, 21(8). <https://doi.org/10.3390/s21082618>
- Wu, X., Sahoo, D., & Hoi, S. C. H. (2020). Recent advances in deep learning for object detection [Online]. Neurocomputing, 396, 39-64. <https://doi.org/10.1016/j.neucom.2020.01.085>
- Yang, D., Wang, X., Zhang, H., Yin, Z. -yu, Su, D., & Xu, J. (2021). A Mask R-CNN based particle identification for quantitative shape evaluation of granular materials [Online]. Powder Technology, 392, 296-305. <https://doi.org/10.1016/j.powtec.2021.07.005>
- Zaridis, D. I., Mylona, E., Tachos, N., Pezoulas, V. C., Grigoriadis, G., Tsiknakis, N., et al. (2023). Region-adaptive magnetic resonance image enhancement for improving CNN-based segmentation of the prostate and prostatic zones [Online]. Scientific Reports, 13(1). <https://doi.org/10.1038/s41598-023-27671-8>
- Zhang, Y., Wang, Z., Song, R., Yan, C., & Qi, Y. (2022). Detection-by-tracking of traffic signs in videos [Online]. Applied Intelligence, 52(7), 8226-8242. <https://doi.org/10.1007/s10489-021-02838-w>

Další bibliografické údaje

Type of thesis: master

Author: Bc. Lukáš BENEDA

Academic Year: 2022/2023

Department: Department of Applied Mathematics and computer science, Faculty of Economics, University of South Bohemia in Ceske Budejovice, Czech Republic

Thesis supervisor: doc. RNDr. Petr Bartoš, Ph.D.

Number of pages: 59

Language of thesis: CZ

Title of Thesis: Segmentation of thermal camera images using selected convolutional neural networks

Annotation: This thesis deals with the issue of instance segmentation of cattle's hoof in thermographic images. The aim of this thesis was to test several solutions and evaluate them. The basis of this thesis is review of existing solutions and state-of-the-art in this field of study, dataset preparation, selection of neural network models and evaluation of the results of individual models. The thesis describes the progress of the work and in the conclusion the individual results are compared and the best solution is evaluated. As results of this thesis are the created dataset of thermographic images of cattle's hoofs and 3 tested and evaluated models of neural networks, from which 2 of them are useable as solution for this issue.

Key words: Thermography, neural nets, cattle's hoof, YOLOv7, Mask R-CNN, Mask DINO, instance segmentation

Přílohy vázané v práci: