Mixing Augmentation and Knowledge-Based Techniques in Unsupervised Domain Adaptation for Segmentation of Edible Insect States

Paweł Majewski[1][a], Piotr Lampa[2][b], Robert Burduk[1][c] and Jacek Reiner[2][d]
[1]*Faculty of Information and Communication Technology, Wrocław University of Science and Technology, Poland*
[2]*Faculty of Mechanical Engineering, Wrocław University of Science and Technology, Poland*

Keywords: Augmentation, Domain Adaptation, Instance Segmentation, Edible Insects, Tenebrio Molitor.

Abstract: Models for detecting edible insect states (live larvae, dead larvae, pupae) are a crucial component of large-scale edible insect monitoring systems. The problem of changing the nature of the data (domain shift) that occurs when implementing the system to new conditions results in a reduction in the effectiveness of previously developed models. Proposing methods for the unsupervised adaptation of models is necessary to reduce the adaptation time of the entire system to new breeding conditions. The study acquired images from three data sources characterized by different types of cameras and illumination and checked the inference quality of the model trained in the source domain on samples from the target domain. A hybrid approach based on mixing augmentation and knowledge-based techniques was proposed to adapt the model. The first stage of the proposed method based on object augmentation and synthetic image generation enabled an increase in average AP_{50} from 58.4 to 62.9. The second stage of the proposed method, based on knowledge-based filtering of target domain objects and synthetic image generation, enabled a further increase in average AP_{50} from 62.9 to 71.8. The strategy of mixing objects from the source domain and the target domain (AP_{50}=71.8) when generating synthetic images proved to be much better than the strategy of using only objects from the target domain (AP_{50}=65.5). The results show the great importance of augmentation and a priori knowledge when adapting models to a new domain.

1 INTRODUCTION

Edible insects are one of the most promising alternative sources of novel food. The number of large-scale edible insect farms is increasing yearly due to the possibility of obtaining a high-protein product at a relatively low-cost (Dobermann et al., 2017). Edible insect breeding is a good solution for utilizing unused areas of livestock buildings where animal diseases such as ASF (African swine fever) previously occurred (Thrastardottir et al., 2021). The need to measure breeding parameters and detect anomalies, combined with the large-scale nature of breeding, necessitates using a dedicated automated monitoring system.

There have recently been few works regarding monitoring edible insect breeding related to the Tenebrio Molitor. (Majewski et al., 2022) proposed a multi-purpose 3-module system, enabling the detection of edible insect growth stages and anomalies (dead larvae, pests), semantic segmentation of feed, chitin, and frass, and larvae phenotyping. The authors used synthetic images generated from a pool of objects, significantly reducing model development time. Other works were based on solutions dedicated to single issues, e.g. classification of larvae segments (Baur et al., 2022), and classification of the gender of pupae (Sumriddetchkajorn et al., 2015). Undoubtedly, the results presented in this works demonstrate the feasibility of using methods based on machine learning and computer vision to inspect edible insect breeding effectively. However, adapting the developed methods to new breeding conditions is still an open problem.

In the literature, we can find a significant number of unsupervised model adaptation methods for the problems of image classification (Madadi et al., 2020), semantic segmentation (Toldo et al., 2020), or object detection (Oza et al., 2021); however, there are fewer works in the area of instance segmentation. Among the most important domain adapta-

[a] https://orcid.org/0000-0001-5076-9107
[b] https://orcid.org/0000-0001-8009-6628
[c] https://orcid.org/0000-0002-3506-6611
[d] https://orcid.org/0000-0003-1662-9762

Majewski, P., Lampa, P., Burduk, R. and Reiner, J.
Mixing Augmentation and Knowledge-Based Techniques in Unsupervised Domain Adaptation for Segmentation of Edible Insect States.
DOI: 10.5220/0011603500003417
In *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2023) - Volume 5: VISAPP*, pages 380-387
ISBN: 978-989-758-634-7; ISSN: 2184-4321

tion methods are discrepancy-based (Csurka et al., 2017; Saito et al., 2018), adversarial-based (including generative-based) (Yoo et al., 2016; Murez et al., 2018), reconstruction-based (including graph-based) (Cai et al., 2019) and self-supervision-based (Khodabandeh et al., 2019; Shin et al., 2020). A relatively simple and intuitive approach to domain adaptation is pseudo-label-based self-training, which involves training the model for the target domain based on samples with pseudo-labels representing a prediction of the model trained on labelled samples from the source domain. An important element in this approach is prediction filtering.

The pseudo-label-based self-training approach seems suitable for instance segmentation and even easier to apply than in object detection. Namely, having masks for objects, it is possible to extract them from images, add them to appropriate object pools and use them further to generate synthetic images. It is also easier to perform filtering at the object level, as it is possible to calculate features for a specific object.

This work proposed a two-stage hybrid method for domain adaptation based on using pseudo-labels for self-training. In 1st stage, it was proposed to expand the training set of samples through augmentation at the image and object levels to reduce the overfitting of the model on the source domain. In 2nd stage, filtering of the obtained predictions was carried out using domain knowledge. An essential contribution of this work is the study of the importance of creating the training set in the 1st and 2nd stages, especially the concept of mixing real and synthetic samples and mixing samples from the source domain (with real labels) and the target domain (with pseudo-labels). In addition, the consequences of using only synthetic data (no real labelled samples in the training set) on the model's performance in cases of inference in and out of the domain were also examined.

2 MATERIAL AND METHODS

2.1 Problem Definition

The problem addressed is detection and segmentation from images of three states of edible insects, namely (1) live larvae, (2) dead larvae, and (3) pupae. The samples are in the form of 512x512 images and come from three sources associated with different types of recording cameras and lighting, namely (1) CA, (2) LU, and (3) JA. Examples of samples from the considered sources, along with the type of objects detected, are shown in Figure 1.

Figure 1: Examples of samples from the considered sources: (a) RGB images, (b) types of detected objects.

The main objective of the research was to propose a suitable domain adaptation method to train the model on one data source (source domain) with labelled samples and make inference on another (target domain) with unlabelled samples with relatively low error. The proposed method is expected to reduce the destructive effect of domain shift on the accuracy of target domain prediction.

2.2 Data Sources

The samples were acquired using three image acquisition systems, differing in the cameras and lighting used. The first one (CA) was an experimental station with a EOS 50D camera (Canon, Tokyo, Japan) with a resolution of 5184 x 3456 pixels and a zoom lens. Diffuse white fluorescence lighting was used. The second (LU) was a data acquisition station purposely built for imaging insects in breeding boxes. It used a Phoenix PHX120S-CC (LUCID Vision Labs, Richmond, Canada) camera with a resolution of 4096 x 3000 pixels and a 12 mm focal length lens. Samples were illuminated with neutral white LEDs in a diffusion tunnel. The third (JA) was a machine vision system prepared for industrial implementation for Tenebrio Molitor breeding. A GOX-12401C-PGE (JAI, Copenhagen, Denmark) camera was used, with a resolution of 4096 x 3000 pixels and a 12 mm lens. In this case, due to size limitations, LED strips providing cold white light were used for direct illumination.

2.3 Dataset

A dataset was prepared for the study, containing samples from various defined sources along with marked object masks from the defined classes. A total of 15 samples from CA, 29 samples from LU and 36 samples from JA were labelled. A summary of the labelled number of objects can be found in Table 1.

Table 1: The number of objects from defined classes in the considered image sources.

source type	object type	no. of objects
CA	live larvae	656
	dead larvae	250
	pupae	124
LU	live larvae	163
	dead larvae	55
	pupae	83
JA	live larvae	1247
	dead larvae	148
	pupae	187

2.4 Data Exploration

For initial data exploration and qualitative evaluation of domain shift, PCA and visualization of selected components were performed. The FID (Frechet Inception Distance) metric (Heusel et al., 2017) was also calculated as a measure of the similarity of features extracted from images belonging to different sources. Lower values of the FID metric mean higher similarity of sample distributions. FID and PCA were based on a feature vector of length 2048 extracted from the last pooling layer of the deep convolutional neural network Inceptionv3 (Szegedy et al., 2015). Masked images of objects (without surrounding background) were used for feature extraction.

2.5 Domain Adaptation with Mixing Augmentation and Knowledge-Based Techniques

The proposed adaptation method consists of two stages described in detail in the following sections. The first stage is based on the augmentation of source domain objects and the generation of synthetic images. The second stage considers filtering target objects based on domain knowledge and re-generating synthetic images using new target domain objects. The idea scheme of the proposed solution is shown in Figure 2.

The method for generating synthetic images involved randomly placing objects on the background image and allowing the simulation of object overlap in dense scenes. Each generated synthetic image was associated with an automatically generated label. The method of generating synthetic images is described in more detail in (Majewski et al., 2022; Toda et al., 2020).

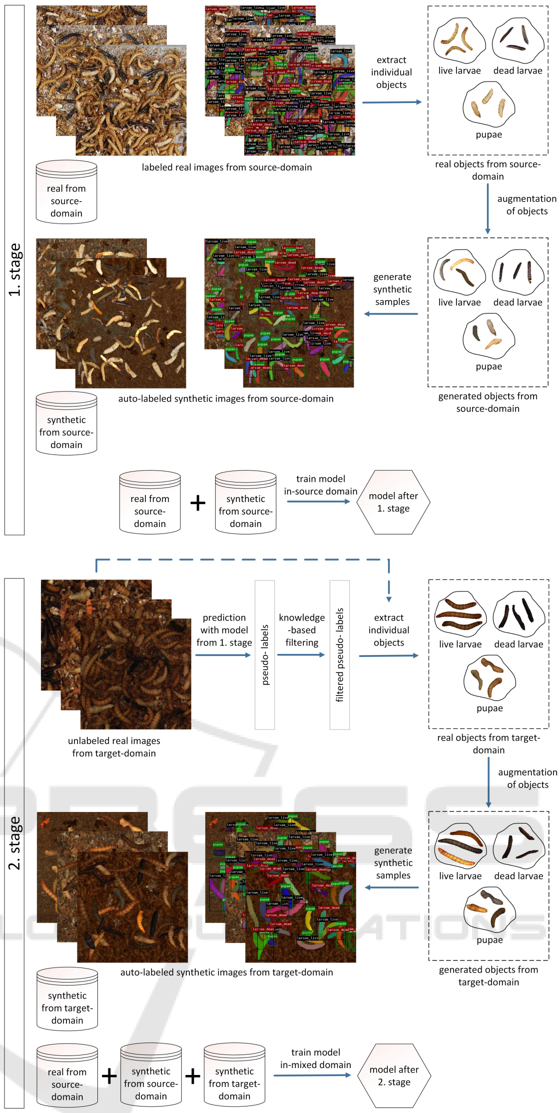

Figure 2: Idea scheme of the proposed solution detailing two stages.

2.5.1 First Stage of Approach with Objects Augmentation

The basis for training models is a set of real labelled samples from the source domain. Evaluation results for a model trained only on a set of real samples (the only_real method) were used as a reference for the following proposed approaches.

Object-level augmentation and synthetic image generation were proposed to extend the source domain samples distribution. First, individual objects were extracted from real images. Then, these objects were augmented, modifying colour, contrast, sharpness and brightness. The generated objects were then placed on the background image, obtaining automatically labelled synthetic images. Examples of the augmented objects and the generated synthetic images are shown in Figure 2.

Three possibilities for constructing the training set for 1st stage were identified. The only_real method assumes training only on real data, the real_synthetic method - on real and synthetic data, and the only_synthetic method - only on synthetic data. For each setting, the Mask-RCNN (He et al., 2017) with backbone ResNet-50 (He et al., 2016) model was trained with default training parameters. An implementation of the Mask R-CNN model from the detectron2 (Wu et al., 2019) library was used in the study. The part related to the 1st stage in Figure 2 shows the real_synthetic approach for creating the training set.

2.5.2 Second Stage of Approach with Knowledge-Based Filtering

The first component of 2nd stage of the proposed solution is an inference using the model trained in 1st stage on unlabelled target domain samples. The resulting predictions were treated as pseudo-labels that needed to be filtered to remove false positive predictions. For filtering, it was used a priori domain invariant knowledge, namely: (1) live larvae are the majority class (see in Table 1), (2) objects from the classes live larvae and dead larvae are the longest (have the largest dimension of the longer side of the bounding box), (3) objects from the dead larvae class have the lowest pixel intensity, (4) objects from the pupae class have the highest pixel intensity. The proposed knowledge-based filtering assumes successively:

1. selection of objects with a minimum length of the longer side of the bounding box d_{min}, with a predicted class live larvae,
2. removal of outliers including mean intensity, size, and length of the longer side of the bounding box among the objects extracted in 1st step, obtaining a distribution of samples representing live larvae,
3. calculation of intensity limits x_{min}, x_{max} for samples representing live larvae,
4. selection of objects with intensity values greater than x_{max}, with predicted class pupae,
5. selection of objects with intensity values less than x_{min}, with predicted class larvae dead.
6. removal of outliers among the objects extracted in the 4th and 5th steps, obtaining a distribution of samples representing pupae and dead larvae.

The obtained new samples in the form of target domain objects and new generated objects after augmentation are then used to generate synthetic data. In 2nd stage, we have available the following sample distributions: (1) real from the source domain, (2) synthetic from the source domain, (3) synthetic from the target domain. The study investigated the following training strategies: the "only target domain samples" strategy assumes training the model only on synthetic data from the target domain, and the "mixed source/target domain samples" strategy assumes mixing samples from the source domain and target domain in the training set. Considering the "mixed source/target domain samples" strategy, in all the approaches identified in 1st stage (only_real, real_synthetic, only_synthetic), the training set defined in 1st stage is extended with synthetic samples from the target domain. Figure 2 shows the "mixed source/target domain samples" strategy with the real_synthetic variant.

2.6 Evaluation

The proposed methods were evaluated using the average precision AP_{50} metric, a standard metric for the evaluation in object detection tasks. The value of the AP_{50} metric represents the area under the precision-recall curve after appropriately interpolating the chart fragments. The AP_{50} metric assumes a threshold value of intersection over union (IoU) 50% between the true and predicted bounding box to consider the prediction significant. Details regarding the calculation of the AP_{50} metric can be found in (Majewski et al., 2022; Padilla et al., 2020).

For the study, 6 possible cases of out-domain crossing (source domain $\rightarrow$ target domain) were selected, namely CA $\rightarrow$ LU, CA $\rightarrow$ JA, LU $\rightarrow$ CA, LU $\rightarrow$ JA, JA $\rightarrow$ CA, JA $\rightarrow$ LU. Evaluation for the out-domain inference cases was carried out for all samples from the target domain. The AP_{50} values for in-domain inference were also determined as a reference. For the in-domain case, the entire dataset was divided into training data (80% of samples) and test data. Evaluation was performed on the test set.

3 RESULTS AND DISCUSSION

As part of the data exploration, PCA was performed, and FID metrics were calculated between sample distributions. A visualization of the selected components for samples from all data sources and defined classes can be found in Figure 3. The calculated FID values can be found in Table 2.

Based on the results from Figure 3 and Table 2, it can be seen that objects from the live larvae class are most similar to each other between distributions.

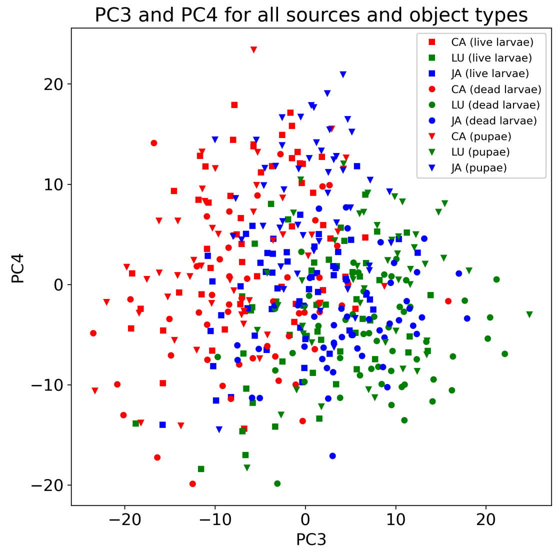

Figure 3: Selected principal components for domain shift exploration based on deep features (Inceptionv3).

Table 2: Comparison of calculated FID metrics between sources based on real samples.

sources	object type	FID
CA and LU	live larvae	124
	dead larvae	166
	pupae	144
	all (average)	145
CA and JA	live larvae	69
	dead larvae	110
	pupae	113
	all (average)	97
LU and JA	live larvae	97
	dead larvae	120
	pupae	115
	all (average)	111

The FID distances between (CA and JA) and (LU and JA) distributions are smaller than the distance between (CA and LU) distributions, which is also confirmed by Figure 3. For the selected components (PC3 and PC4), samples from JA (blue markers) are between samples from CA (red) and LU (green).

A comparison of different domain adaptation methods can be found in Table 3 for the "mixed source/target domain samples" strategy and in Table 4 for the "only target domain samples" strategy. As reference values for assessing the quality of domain adaptation are the results obtained by the models trained and tested in-domain presented in Table 5.

The results obtained for 1st stage of model adaptation (Table 3) prove that the real_synthetic method (average $AP_{50} = 62.9$), which assumes the use of both real and synthetic samples for training, is the most suitable for use in the problem under consideration. The use of only synthetic samples (only_synthetic method, average $AP_{50} = 54.4$) or only real samples (only_real method, average $AP_{50} = 58.4$) may be better in special cases (only_synthetic for LU $\rightarrow$ CA, only_real for LU $\rightarrow$ JA, CA $\rightarrow$ JA), but in general (averaged), these approaches achieve smaller AP values than the real_synthetic method. For the special cases mentioned above, the difference between the best-obtained result and the AP value for the real_synthetic method did not exceed $\Delta AP_{50} = 4.0$. On the other hand, for the LU $\rightarrow$ CA, the difference between the AP values for only_real and real_synthetic was $\Delta AP_{50} = 14.0$, and for the JA $\rightarrow$ CA was $\Delta AP_{50} = 9.4$, which is a significant difference in the effectiveness of the models.

Using only synthetic data for model training can significantly speed up the process of developing models (Majewski et al., 2022); however, based on the results obtained in this research, we can observe that this is associated with the risk of losing model accuracy. This observation is confirmed by the results after the 1st and 2nd stages of domain adaptation for inference out-domain in Table 3 (the only_synthetic approach was characterized by $\Delta AP_{50} = 8.5$ lower AP_{50} than the real_synthetic approach in the 1st stage and by $\Delta AP_{50} = 4.4$ lower AP_{50} in the 2nd second). The lack of real data in the training set mostly affects the results for in-domain inference in Table 5 ($\Delta AP_{50} = 11.3$ difference between only_synthetic and real_synthetic approaches).

When considering the results separately for each of the defined classes, it should be noted that objects of the live larvae class are the easiest to detect (average AP_{50} after 2nd stage – 81.8) after a domain change, while objects of the pupae class – the most difficult (average AP_{50} after 2nd stage – 66.6). This is consistent with initial conclusions from data exploration based on FID values in Table 2.

Quantitative indicators confirm the importance of augmentation in 1st stage for the real_synthetic approach. Additional samples complement the relevant places in the feature space and can expand the distribution of features for a given class.

Analyzing the results from 2nd stage for the two proposed strategies in Table 3 for "mixed source/target samples" strategy and in Table 4 for "only target samples" strategy, we can conclude that the "mixed source/target samples" strategy is the most suitable for creating a training set, which is confirmed by obtaining an increased AP_{50} from 65.5 to 71.8 compared to the "only target samples" strategy.

A summary of the most important results ob-

Table 3: Comparison of proposed domain adaptation methods for mixed source/target domain samples strategy.

case type	method	AP_{50}							
		stage 1.				stage 2. (mixing strategy)			
		live l.	dead l.	pup.	avg.	live l.	dead l.	pup.	avg.
CA → LU	only_real	64.9	58.9	74.7	66.2	79.9	59.7	75.4	71.7
	real_synthetic	70.7	61.0	78.0	**69.9**	82.8	61.7	78.1	**74.2**
	only_synthetic	63.3	54.0	65.5	60.9	82.1	62.0	77.4	73.8
CA → JA	only_real	69.7	50.4	29.2	**49.8**	75.3	55.1	36.7	55.7
	real_synthetic	72.2	38.5	27.6	46.1	76.0	55.0	36.5	**55.8**
	only_synthetic	59.3	28.6	18.6	35.5	77.3	40.8	31.7	49.9
LU → CA	only_real	41.3	58.5	39.5	46.4	79.7	78.1	69.4	**75.7**
	real_synthetic	65.1	68.8	47.2	60.4	80.2	76.4	69.6	75.4
	only_synthetic	64.3	69.3	49.9	**61.2**	79.8	76.7	70.6	**75.7**
LU → JA	only_real	73.8	53.8	28.7	**52.1**	83.3	66.9	56.2	68.8
	real_synthetic	74.1	45.9	26.2	48.7	84.9	62.3	62.9	**70.0**
	only_synthetic	59.3	31.6	14.7	35.2	83.2	50.5	55.2	63.0
JA → CA	only_real	76.8	67.1	47.8	63.9	84.6	76.4	66.8	75.9
	real_synthetic	78.8	73.3	67.7	**73.3**	82.9	76.3	69.5	**76.2**
	only_synthetic	71.4	73.6	61.1	68.7	78.5	72.8	59.8	70.4
JA → LU	only_real	75.2	68.3	71.5	71.7	83.2	74.3	84.2	**80.6**
	real_synthetic	82.9	71.5	82.7	**79.0**	84.2	70.6	82.8	79.2
	only_synthetic	74.2	60.3	60.3	64.9	79.6	61.9	73.3	71.6
all (summary)	only_real	67.0	59.5	48.6	58.4	81.0	68.4	64.8	71.4
	real_synthetic	74.0	59.8	54.9	**62.9**	81.8	67.1	66.6	**71.8**
	only_synthetic	65.3	52.9	45.0	54.4	80.1	60.8	61.3	67.4

Table 4: Results for the only target samples strategy for the 2nd stage of domain adaptation.

case type	method	AP_{50}			
		stage 2. (only target samples strategy)			
		live larvae	dead larvae}	pupae	average
all (summary)	only_real	76.6	60.9	55.8	64.5
	real_synthetic	78.5	59.1	59.0	**65.5**
	only_synthetic	77.5	54.7	54.2	62.2

Table 5: Reference values for domain adaptation as AP_{50} values for in-domain inference.

source type	method	AP_{50}			
		live larvae	dead larvae	pupae	average
all (summary)	only_real	86.6	78.8	84.4	83.3
	real_synthetic	88.4	81.8	85.2	**85.2**
	only_synthetic	80.0	71.6	70.3	73.9

tained in the study is presented on the radar plot in Figure 4. In Figure 4 it can be seen that for the crossings JA → CA, JA → LU, CA → LU, already 1st stage of the proposed method based on augmentation achieves reasonable AP results when using the real_synthetic approach. The 2nd stage caused a significant increase in AP for the crossings LU → JA and LU → CA. After the two stages of the proposed solution, the final value of the obtained AP values strongly depended on the target domain. For crossings where the target domain was JA, the final AP values were the lowest (AP_{50} =55.8, AP_{50} =70). In summary, the best variation of the proposed method made it possible to increase the average AP_{50} from 58.4 to 62.9 after 1st stage and to 71.8 after the 2nd stage. To obtain as high AP_{50} values as in-domain trained models (AP_{50} = 85.2), additional labelling should be performed, especially of objects undetected by models after the 2nd stage of adaptation. The obtained AP_{50} level between 70 and 80.6 for 5 out of 6 (except for CA → JA) types of crossings between domains makes it possible to improve additional labelling by label proposals that are predictions of the model obtained after the 2nd stage.

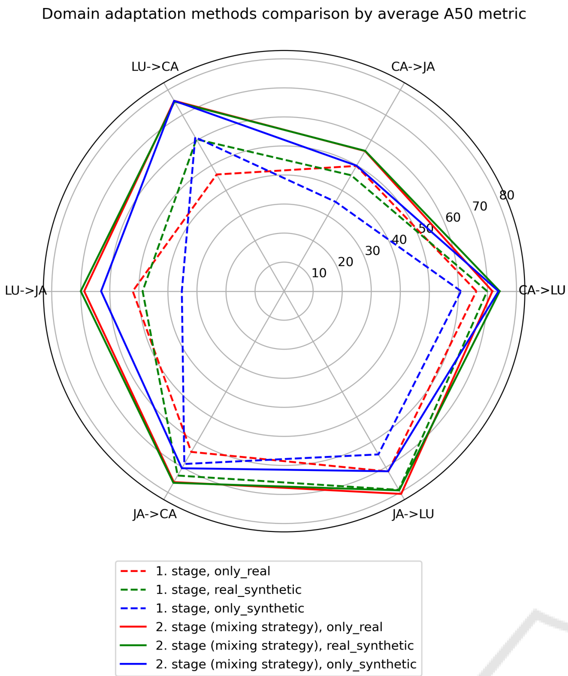

Figure 4: Comparison of proposed domain adaptation methods for different cases.

To confirm the good quality of predictions after domain adaptation, Figure 5 compares the predictions by the in-domain trained model with the predictions of the model after domain adaptation for three selected samples from different target domains. For clarity, Figure 5 shows the predictions only for the dead larvae and pupae classes.

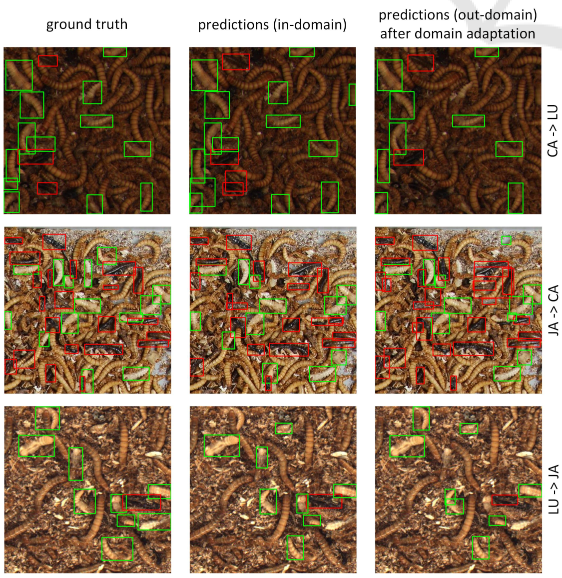

Figure 5: Comparison of predictions with ground truth for in-domain and out-domain inference cases.

4 CONCLUSIONS

The proposed two-stage method for domain adaptation made it possible to significantly increase the efficiency of object detection (AP_{50} increased from 58.4 to 71.8) when changing the domain without additional user supervision. The best results were obtained when the final training set consisted of real samples from the source domain, synthetic samples from the source domain and synthetic samples from the target domain (associated with filtered objects from the target domain). It confirms the validity of mixing real and synthetic samples in the training set and mixing objects from the source and target domains. It can also be concluded from the results that using only synthetic data when training models can significantly reduce the efficiency of the models for both in-domain and out-domain inference. The study showed the importance of augmentation techniques and consideration of a priori knowledge for domain adaptation.

The proposed method is flexible and can be extended to other classes of objects representing states of edible insects, e.g., beetles. The method's extension would include adding new rules when filtering the prediction based on a priori knowledge. The developed solutions will undoubtedly help rapidly adapt monitoring systems for breeding the Tenebrio Molitor to new breeding conditions.

Future research should focus on increasing the quality of synthetic data. An interesting research direction is to develop synthetic images based only on a priori knowledge independently of a specific domain. This approach could obtain a basic model not overfitted on a particular domain.

ACKNOWLEDGEMENTS

We wish to thank Mariusz Mrzygłód for developing applications for the designed data acquisition workstation. We wish to thank Paweł Górzyński and Dawid Biedrzycki from Tenebria (Lubawa, Poland) for providing a data source of boxes with Tenebrio Molitor. The work presented in this publication was carried out within the project “Automatic mealworm breeding system with the development of feeding technology” under Sub-measure 1.1.1 of the Smart Growth Operational Program 2014-2020 co-financed from the European Regional Development Fund on the basis of a co-financing agreement concluded with the National Center for Research and Development (NCBiR, Poland); grant POIR.01.01.01-00-0903/20.

REFERENCES

Baur, A., Koch, D., Gatternig, B., and Delgado, A. (2022). Noninvasive monitoring system for tenebrio molitor larvae based on image processing with a watershed algorithm and a neural net approach. *Journal of Insects as Food and Feed*, pages 1–8.

Cai, Q., Pan, Y., Ngo, C.-W., Tian, X., Duan, L., and Yao, T. (2019). Exploring object relation in mean teacher for cross-domain detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11457–11466.

Csurka, G., Baradel, F., Chidlovskii, B., and Clinchant, S. (2017). Discrepancy-based networks for unsupervised domain adaptation: a comparative study. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 2630–2636.

Dobermann, D., Swift, J., and Field, L. (2017). Opportunities and hurdles of edible insects for food and feed. *Nutrition Bulletin*, 42(4):293–308.

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30.

Khodabandeh, M., Vahdat, A., Ranjbar, M., and Macready, W. G. (2019). A robust learning approach to domain adaptive object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 480–490.

Madadi, Y., Seydi, V., Nasrollahi, K., Hosseini, R., and Moeslund, T. B. (2020). Deep visual unsupervised domain adaptation for classification tasks: a survey. *IET Image Processing*, 14(14):3283–3299.

Majewski, P., Zapotoczny, P., Lampa, P., Burduk, R., and Reiner, J. (2022). Multipurpose monitoring system for edible insect breeding based on machine learning. *Scientific Reports*, 12(1):1–15.

Murez, Z., Kolouri, S., Kriegman, D., Ramamoorthi, R., and Kim, K. (2018). Image to image translation for domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4500–4509.

Oza, P., Sindagi, V. A., VS, V., and Patel, V. M. (2021). Unsupervised domain adaptation of object detectors: A survey. *arXiv preprint arXiv:2105.13502*.

Padilla, R., Netto, S. L., and Da Silva, E. A. (2020). A survey on performance metrics for object-detection algorithms. In *2020 international conference on systems, signals and image processing (IWSSIP)*, pages 237–242. IEEE.

Saito, K., Watanabe, K., Ushiku, Y., and Harada, T. (2018). Maximum classifier discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3723–3732.

Shin, I., Woo, S., Pan, F., and Kweon, I. S. (2020). Two-phase pseudo label densification for self-training based domain adaptation. In *European conference on computer vision*, pages 532–548. Springer.

Sumriddetchkajorn, S., Kamtongdee, C., and Chanhorm, S. (2015). Fault-tolerant optical-penetration-based silkworm gender identification. *Computers and Electronics in Agriculture*, 119:201–208.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.

Thrastardottir, R., Olafsdottir, H. T., and Thorarinsdottir, R. I. (2021). Yellow mealworm and black soldier fly larvae for feed and food production in europe, with emphasis on iceland. *Foods*, 10(11):2744.

Toda, Y., Okura, F., Ito, J., Okada, S., Kinoshita, T., Tsuji, H., and Saisho, D. (2020). Training instance segmentation neural network with synthetic datasets for crop seed phenotyping. *Communications biology*, 3(1):1–12.

Toldo, M., Maracani, A., Michieli, U., and Zanuttigh, P. (2020). Unsupervised domain adaptation in semantic segmentation: a review. *Technologies*, 8(2):35.

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., and Girshick, R. (2019). Detectron2. https://github.com/facebookresearch/detectron2.

Yoo, D., Kim, N., Park, S., Paek, A. S., and Kweon, I. S. (2016). Pixel-level domain transfer. In *European conference on computer vision*, pages 517–532. Springer.