Continuous Image Classification on Data Streams using Contrastive Learning and Cluster Analysis

Andreas Schliebitz[1]

Abstract: This PhD proposal presents a concept for an AI-based computer vision system trained on image data streams for real-time classification of unlabeled objects. The required AI models are trained through an underexplored combination of self-supervised and incremental learning. Emphasis will be placed on contrastive learning using the SimCLR framework and its successors. The need for such a system is motivated by the observation that supervised learning approaches often require large labeled datasets. The labeling process, which is usually performed manually, is not only time-consuming but inherently prone to errors. For sufficiently large image data streams, timely labeling of samples becomes impossible leading to sporadic data annotation cycles and the capture of only temporarily representative features. Such an approach might also render the resulting classifier vulnerable to domain shift and concept drift. The image data stream used in this proposal consists of unlabeled color images of clean potatoes, which are to be sorted into several defect classes by a self-supervised classifier. Contrastive transfer learning is performed on this image data stream for the selection of a feature extractor. In this approach, different pre-trained backbone architectures are adapted and evaluated using the SimCLR framework. The classifiers are evaluated based on their generated feature vectors using cluster analysis. This involves searching for novel evaluation methods that do not require labels and are more suitable for judging model performance than existing methods. Furthermore, by clustering the feature vectors, an automatic and adaptive classification might be achievable without the use of labels. In a subsequent step, the self-supervised classifiers are continuously improved using incremental learning methods. For this the models are incrementally trained on the image data stream over a longer period of time. Potential adjustments to the data stream could increase the classifier's accuracy as well as make it more robust to domain adaptation problems. A final validation of the incrementally self-learning classification system can be performed with smaller, manually annotated datasets.

Keywords: Feature Learning, Self-Supervised Learning, Contrastive Learning, Incremental Learning, Image Classification, Data Stream Processing, Cluster Analysis, Computer Vision

1 Problem Statement

In many practical use cases of artificial intelligence (AI), it is observed that the training of AI models is often hindered by a lack of labeled data [Al23]. AI developers predominantly encounter these problems in the context of supervised learning, a widely used machine learning (ML) technique [KG20]. The supervised learning approach uses an existing dataset composed of individual observations, also known as samples. These samples adopt different

[1] Osnabrück University of Applied Sciences, Faculty of Engineering and Computer Science (IuI), Albrechtstr. 30, 49076 Osnabrück, Germany, a.schliebitz@hs-osnabrueck.de

doi:10.18420/ki2023-dc-04

shapes in the various machine learning disciplines. In the specific case of image classification, which this paper will focus on, a dataset consists of images depicting objects belonging to different classes. In supervised learning, these objects must be labeled before a classification model can be trained [SS18]. Given a classification task involving n possible classes, a label can be conceived as a natural number l with $1 \leq l \leq n$, which is assigned to each object present in the dataset. This assignment is usually carried out during the labeling phase by humans using a label's textual representation, providing the AI with its learning objective. During the training phase, the AI uses these labeled inputs to learn a function that ideally produces the correct label for all subsequent inputs outside the training dataset.

There are applications in research as well as in industry where the preparation of a labeled dataset has to be performed manually. In many cases, the reason for this is the expert knowledge required for correctly assigning class labels to objects. Labeling is an error-prone but nonetheless important process for building powerful AI models [Wh09]. Contamination of a dataset with incorrectly labeled objects can distort the output of the resulting AI model, reducing its value to researchers and industry professionals [NDG20].

Today, the labeling effort can already be reduced with auto-annotation techniques. However, this often happens at the expense of label quality [Al21]. Auto-annotation approaches usually involve AI models that have been pre-trained on large labeled datasets such as ImageNet or COCO. These generic models can therefore only be applied with limitations to application-specific datasets, often leading to unsatisfactory labeling results. This is especially the case in the context of class labeling, since a class assignment can only be performed correctly or incorrectly. This is in contrast to object detection and image segmentation, where slight variations in bounding box or segmentation mask accuracy hardly affect the usefulness of the resulting model. Apart from the fact that in some use cases the training data has to be annotated by humans due to quality assurance reasons, there are even more extreme scenarios where a lot of training data can be accumulated in a short period of time. If such a data stream exceeds a certain size, providing sufficient human resources is often not possible.

The PhD project presented in this paper is motivated by a use case within the food industry facing the exact problems described above. As part of the research project Agri-Gaia, the Osnabrück University of Applied Sciences is working together with the Wernsing Feinkost GmbH (Wernsing) in order to develop a computer vision system for the quality assessment of potatoes [Sc23]. This system is expected to locate potatoes on a conveyor belt and sort them into at least one of eleven defect classes. Since a potato can exhibit multiple defects, this classification task is treated as a multi-label problem. The aim is to record the quality distribution of a potato delivery and to replace the currently sample-based inspection. Wernsing receives about four million potatoes a day, which are individually recorded using a camera system developed as part of the research project. Neither the Osnabrück University of Applied Sciences nor Wernsing are in a position to annotate these data quantities manually. Therefore, only a tiny fraction of the features contained in the millions of individual images can be used for the supervised training of a defect classification model.

2 Objectives

According to the problems explained in the previous section, an alternative approach is needed that allows for both the training and the continuous adaptation of a classification model without labeled image data. The refinement of the classifier using recently recorded samples has to be incremental, since repeated training over the entire dataset would take too much time and computational resources.

Since the intended PhD project has considerable relevance especially for science but also for industrial applications, the theoretical concepts presented in this proposal should be implemented in a real-world system. In doing so, a modular, replicable and unintrusive end-to-end computer vision system has to be embedded into already existing processes. Therefore, a practical goal of this PhD project is to design and implement a process for an incremental classification of image data streams using self-learning[2] AI. The development and testing of this concept can be exemplified by the defect classification of potatoes in Wernsing's goods receiving department. The image data stream is generated by a continuous recording of potatoes on a conveyor belt using RGBD cameras. The theoretical objectives of this PhD proposal are mainly based on two factors, which significantly influence the success of image classification without labeled training data: Self-supervised feature extraction and cluster analysis of feature vectors.

First, the selection of the neural network responsible for feature extraction plays a key role in classification. Such a neural network is also called a feature extractor or backbone. In the context of image data, these feature extractors are often classical convolutional neural networks (CNNs) or more advanced vision transformer (ViT) architectures. A feature extractor can be viewed as a function that generates for each input image a feature vector in the associated feature space. The input is said to be embedded in the feature space by its feature vector. The literature contains a number of different feature extractors that are trained on large labeled datasets in an often supervised manner [To23]. These pre-trained models are commonly adapted to an application-specific and typically smaller dataset using supervised transfer learning techniques [PY10]. However, classical transfer learning is unsuitable for this PhD project, since the classification of an image data stream should be accomplished without labels. Therefore, this project aims at investigating the impact of contrastive methods on transfer learning. This self-supervised approach does not require labeled training data, unlike supervised transfer learning. Initially, supervised pre-trained feature extractors will be trained on an image data stream using contrastive learning. Contrastive learning [BCV13] is a form of self-supervised learning (SSL) [Ja20] where the loss function's objective is to decrease the distance between similar samples while simultaneously increasing it to others. Subsequently, both supervised and self-supervised trained feature extractors are evaluated and compared on a small ground truth dataset. Since neither approach requires manual labeling of training data, human effort is comparably low. The results of this experiment will

[2] In the context of this proposal, self-learning AI refers to an AI model that can be successfully trained using appropriate learning techniques and new data without human interaction.

be analyzed to determine if the generalization ability of the SSL-based models outperform their supervised pre-trained counterparts, and if so, to what extent. These findings can also be used to investigate whether the best supervised feature extractors are generally better suited for SSL-based transfer learning.

Second, an effective clustering of the feature vectors in feature space is a central component of label-free classification. Such classification problems can be viewed as a clustering problem in feature space, if the loss function used minimizes the distances between embeddings of similar inputs (like in contrastive learning). A clustering of feature vectors can be used for classification of new inputs by assigning a fixed natural number $l \in \mathbb{N}$ (see sec. 1) to each of the clusters. During inference, the class of an input image is given by the number of the cluster with the smallest distance to its embedding in feature space. The formation of clusters is based on the assumption that similar inputs often belong to the same class. Thus, after contrastively learning their representations, the resulting embeddings should tend to form separable clusters in feature space. Based on this hypothesis, the method used for clustering the feature vectors can also have an impact on the classification result.

Based on these observations, this PhD project intends to examine methods of unsupervised feature learning and apply them to the clustering of feature vectors. The feature vectors will be generated by an AI-based image classifier, which will be trained using self-supervised learning on an image data stream without labels. In the literature, this initial training is often followed by a downstream task that adapts the feature extractor to the actual classification problem using a few labeled samples (see transfer learning). However, research suggests that labeling such a dataset once may be insufficient for image data streams, as new classes may be added over time or existing classes could change in their visual appearance (see domain shift [SFS16]). Automatic clustering of feature vectors is a conceivable approach to detect the emergence of new classes in feature space. It is further worth investigating whether cluster analysis can be used to quantify the quality of an SSL-based model more accurately than existing methods. This includes logistic regression, which is one of the better known evaluation methods for self-supervised models, along with data efficiency and the aforementioned transfer learning. A disadvantage of this comparatively simple evaluation method is that it only evaluates the linear separability of the learned representations.

In addition to a suitable learning technique, an approach for incrementally adapting the classifier to changes in the data stream has to be developed. In the literature, these methods are often studied under the terms of online and incremental learning. From a conceptual point of view, pure online learning is preferable to incremental learning because in online learning the AI model is strictly trained with the latest sample, eliminating the need for managing a dataset. However, pure online learning is hardly suitable in real-world applications, since AI models trained with this technique suffer from catastrophic interference [MC89]. Under this phenomenon, the AI model loses its generalization ability as new samples displace the already learned representations over time. In contrast to online learning, methods which are being developed in the field of incremental learning specifically counteract this problem. Examples of such mechanisms include the incorporation of older samples into the training process and the use of stable incremental learning methods such as Fuzzy ART [CGR91] or TopoART [Ts10].

In summary, the main goal of this PhD project is to combine feature learning approaches (e. g., self-supervised learning and cluster analysis) with incremental learning methods to create a novel self-learning AI system that can classify objects in image data streams without labels.

3 Related Work

With respect to the classification of image data streams, Lima et al. [Li22] summarize numerous problems related to labeled ground truth data. The authors draw attention to the concept of latency, which describes the temporal availability of labels after the acquisition of new samples. Furthermore, they discuss the problems of concept drift and the emergence of new classes in the image data stream, leading to an open set classification problem. They also emphasize that due to slowly evolving image data streams, a classifier should be incrementally adaptable to those changes. These observations are central to research on class-incremental learning [Ma22]. Like in this proposal, feature extraction using convolutional neural networks is considered powerful by the authors. According to Guérin et al. [Gu17], CNNs are also suitable for unsupervised classification tasks. By leveraging the feature extraction capabilities of CNNs, input images are usually represented through their numerical feature vectors. Prototype-based algorithms such as NCM (Nearest Class Mean) or iCaRL (Incremental Classifier and Representation Learning) [Re17] are proposed for an incremental image stream classification. These allow for a dynamic extension of the classifier to include new classes without having to recompute the underlying model. A selection of highly suitable training inputs from a continuous image data stream can be performed using active learning. Novel approaches to feature extraction with pre-trained vision transformers or particularly deep or wide convolutional neural networks are not discussed by Lima et al. Furthermore, no reference to modern feature learning approaches are made, which are represented in this proposal through the SimCLR framework, contrastive learning and cluster analysis.

Adaptation of the classifier to the data stream is achieved by Zheng et al. [ZW22] using an Online-CNN architecture (OCNN). The OCNN architecture is based on the five convolutional layers of AlexNet and uses Hedge Backpropagation (HBP) to adjust the network weights online. In this approach, each convolutional layer is conceived as a separate classifier of different depth. Backpropagation of the prediction error is performed by each of these classifiers, which is expected to have a positive impact on the models generalization ability in online applications. In their approach, the authors assume that the labels become known immediately after new samples are recorded. However, this assumption is not realistic in many real-world scenarios, thus requiring an alternative approach like contrastive transfer learning without labels. In contrast to Lu et al. [LJH21], who only adapt the ResNet50 V2 backbone of the Big Transfer Model (BiT) in a contrastive learning approach using

SimSiam, this PhD project aims at experimenting with a much larger number of recent feature extraction architectures.

The embeddings generated by a feature extractor can be of high dimensionality depending on the architecture. In order to cluster these feature vectors efficiently, the algorithm of Rahman et al. can be used [RJ18]. More generic methods have also been used in the literature for feature clustering. These include HDBSCAN, PCA, and the k-means algorithm [KK22, Ce09]. The result of a clustering-based classification procedure can be evaluated without labels using cluster validation algorithms. A widely used quality metric for convex clusters is Rousseeuw's silhouette coefficient [Ro87]. Once a clustering result contains non-convex clusters, an evaluation can be performed via a Density-Based Clustering Validation (DBCV) given by Moulavi et al. [Mo14].

4 Preliminary Work

Most of the preliminary work for this PhD project was carried out as part of the Agri-Gaia research project at Osnabrück University of Applied Sciences. At the time of this proposal, a camera setup already exists in Wernsing's incoming goods department, consisting of three RGBD cameras, a lighting setup and an edge computer. The cameras are mounted side by side in an aluminum frame and aligned orthogonally to a two-meter wide conveyor belt. Potatoes coming out of the washing line are photographed from a height of about 57 cm with a resolution of 4032×3040 px. The system operates at a frame rate of about three frames per second, which is derived from the conveyor belt speed and the image height in world coordinates. Using this image data stream as well as semi-synthetically generated training data, a YOLOv5 object detector was trained in a supervised way. The accuracy of this detector is about 0.96 mAP@[.5:.95] and therefore comparable to the performance of a human. Applying this potato detector to Wernsing's incoming potato stream yielded an unannotated RGB dataset of 21 million potatoes with a total size of one terabyte in a period of five days. This dataset can be used in the PhD project to simulate data streams and train different AI models in a self-supervised manner. In addition to this unlabeled dataset, a much smaller labeled dataset exists that can be used as ground truth for evaluating AI-based classification models. This and several other preliminary efforts, such as the supervised training of an EfficientNetV2 classifier, can be found in the publication [Sc23].

5 Approach and Methods

For the training of a first SSL-based classifier, an image dataset without labels is needed first. This is collected in the incoming goods department of Wernsing using the camera setup and object detector described in [Sc23]. It is important to note that the individual samples should be captured over a longer period of time in order to record seasonal changes in the appearance of the potatoes.

Using this dataset, a backbone architecture for feature extraction will be selected in the next step. The weights of this neural network are initially pre-trained and will be adapted to the unlabeled potato dataset using contrastive transfer learning. For this purpose, the default ResNet50 backbone of SimCLR [Ch20] is replaced by compatible feature extractors from the Torchvision library. The weights of these AI models were trained in a classical supervised learning approach on the ImageNet1K dataset [De09]. During the training, InfoNCE (Noise-Contrastive Estimation) [OLV18] is used as the loss function. After all training runs are completed, the effectiveness of the contrastive transfer learning approach is quantified by comparing the SSL-based models to the supervised pre-trained classifiers. The implementation is carried out using the Python programming language with the help of the PyTorch Lightning library.

The evaluation of the SSL-based classification models will be carried out in two different ways. First, the classical approach of a logistic regression using a decoupled linear layer will be implemented. Since this approach only provides an estimate of the linear separability of the learned representations [Re22], cluster analysis techniques will be used to find more meaningful evaluation procedures for SSL-based image classifiers. In contrast to logistic regression, these novel evaluation methods are intended to work without labeled data. The practical implementation of this cluster-based evaluation can be achieved, for example, with the clustering module of the Python library Scikit-learn.

A continuous adaptation of the SSL-based classifier is achieved through a combination of contrastive and incremental learning. For this purpose, the training and evaluation pipelines from previous experiments are reused and integrated into a software framework that enables incremental training of AI models. The image data stream from Wernsing's incoming goods department will serve as the data source. This image data stream can either be used in real time or, for reasons of reproducibility, simulated by a dataset recorded in advance with the same camera setup. Based on the accuracies of the contrastively trained classifiers, the top n backbone architectures are integrated into ContinualAI's Avalanche framework [Lo21]. These backbones are then trained incrementally, evaluated, and compared over a fixed but arbitrary time period on the image data stream. The goal of this comparison is to draw conclusions about the suitability of certain backbone architectures for incremental learning. Within these longer train periods, the accuracy of the classifiers will be recorded at fixed intervals. Based on these observations, negative effects of domain shift, concept drift or catastrophic interference shall be detected at an early stage. By introducing targeted countermeasures, their effectiveness can be investigated experimentally at the same time.

6 Planned Evaluation

After completion of the potato classification system prototyped in Figure 1, its accuracy can be validated using batches of manually labeled data. These ground truth datasets can be created in cooperation with Wernsing and passed through the camera system in the incoming goods department. Reliability checks like these would be carried out at regular intervals to rule out problems linked to domain adaptation. Since potatoes are considered a perishable commodity, manual sampling would need to be performed again prior to each validation cycle. However, due to the incremental learning approach, there is a possibility that the SSL-based classifier will become robust towards seasonally recurring variations in the data stream. Once the model has established this property, manual validation of the classification results could be rendered unnecessary. This would most likely be accompanied by a consolidation of the clusters in feature space, leading to a final replacement of the cluster numbers with their corresponding labels.

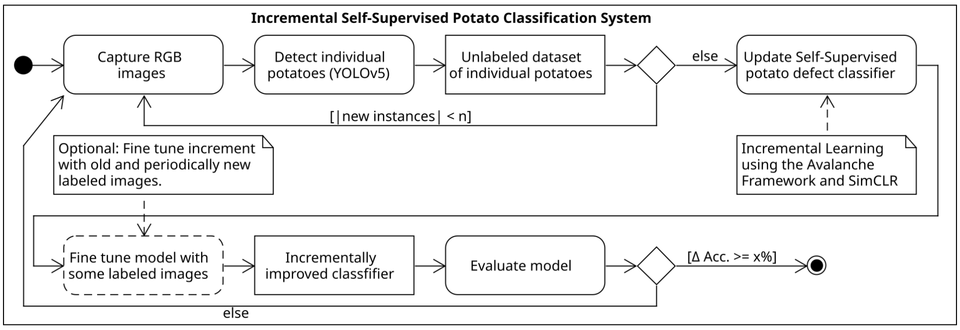

Fig. 1: A potentially feasible train and evaluation cycle of an incrementally learning computer vision system for a self-supervised classification of potato defects.

Acknowledgments

The preliminary work presented in this proposal was funded by the German Federal Ministry of Economics and Climate Protection as part of the research project Agri-Gaia under grant number 01MK21004G.

Bibliography

[Al21] Alhazmi, Khaled; Alsumari, Walaa; Seppo, Indrek; Podkuiko, Lara; Simon, Martin: Effects of annotation quality on model performance. In: 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE, pp. 063–067, 2021.

[Al23] Alzubaidi, Laith; Bai, Jinshuai; Al-Sabaawi, Aiman; Santamaría, Jose; Albahri, A. S.; Al-dabbagh, Bashar Sami Nayyef; Fadhel, Mohammed A.; Manoufali, Mohamed; Zhang, Jinglan; Al-Timemy, Ali H. et al.: A survey on deep learning tools dealing with data scarcity: definitions, challenges, solutions, tips, and applications. Journal of Big Data, 2023.

[BCV13] Bengio, Yoshua; Courville, Aaron; Vincent, Pascal: Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8):1798–1828, 2013.

[Ce09] Celik, Turgay: Unsupervised change detection in satellite images using principal component analysis and k-means clustering. IEEE geoscience and remote sensing letters, 6(4):772–776, 2009.

[CGR91] Carpenter, Gail A.; Grossberg, Stephen; Rosen, David B.: Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system. Neural networks, 4(6):759–771, 1991.

[Ch20] Chen, Ting; Kornblith, Simon; Norouzi, Mohammad; Hinton, Geoffrey: A simple framework for contrastive learning of visual representations. In: International conference on machine learning. PMLR, pp. 1597–1607, 2020.

[De09] Deng, Jia; Dong, Wei; Socher, Richard; Li, Li-Jia; Li, Kai; Fei-Fei, Li: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. Ieee, pp. 248–255, 2009.

[Gu17] Guérin, Joris; Gibaru, Olivier; Thiery, Stéphane; Nyiri, Eric: CNN features are also great at unsupervised classification. arXiv preprint arXiv:1707.01700, 2017.

[Ja20] Jaiswal, Ashish; Babu, Ashwin Ramesh; Zadeh, Mohammad Zaki; Banerjee, Debapriya; Makedon, Fillia: A survey on contrastive self-supervised learning. Technologies, 9(1):2, 2020.

[KG20] Kour, Herleen; Gondhi, Naveen: Machine Learning Techniques: A Survey. In (Raj, Jennifer S.; Bashar, Abul; Ramson, S. R. Jino, eds): Innovative Data Communication Technologies and Application. Springer International Publishing, Cham, pp. 266–275, 2020.

[KK22] Kim, Jiyeon; Kang, Youngok: Automatic classification of photos by tourist attractions using deep learning model and image feature vector clustering. ISPRS International Journal of Geo-Information, 11(4):245, 2022.

[Li22] de Lima, Mateus C.; Souza, YanStivalettie; Faria, Elaine R.; Barioni, Maria Camila N.: A comprehensive analysis of the diverse aspects inherent to image data stream classification. Knowledge and Information Systems, 64(8):2215–2238, 2022.

[LJH21] Lu, Yuzhe; Jha, Aadarsh; Huo, Yuankai: Contrastive learning meets transfer learning: a case study in medical image analysis. arXiv preprint arXiv:2103.03166, 2021.

[Lo21] Lomonaco, Vincenzo; Pellegrini, Lorenzo; Cossu, Andrea; Carta, Antonio; Graffieti, Gabriele; Hayes, Tyler L.; De Lange, Matthias; Masana, Marc; Pomponi, Jary; Van de Ven, Gido M. et al.: Avalanche: an end-to-end library for continual learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3600–3610, 2021.

[Ma22] Masana, Marc; Liu, Xialei; Twardowski, Bartłomiej; Menta, Mikel; Bagdanov, Andrew D.; Van De Weijer, Joost: Class-incremental learning: survey and performance evaluation on image classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(5):5513–5533, 2022.

[MC89] McCloskey, Michael; Cohen, Neal J.: Catastrophic interference in connectionist networks: The sequential learning problem. In: Psychology of learning and motivation, volume 24, pp. 109–165. Elsevier, 1989.

[Mo14] Moulavi, Davoud; Jaskowiak, Pablo A.; Campello, Ricardo J. G. B.; Zimek, Arthur; Sander, Jörg: Density-based clustering validation. In: Proceedings of the 2014 SIAM international conference on data mining. SIAM, pp. 839–847, 2014.

[NDG20] Nigam, Nitika; Dutta, Tanima; Gupta, Hari Prabhat: Impact of noisy labels in learning techniques: A survey. In: Advances in Data and Information Sciences: Proceedings of ICDIS 2019. Springer, pp. 403–411, 2020.

[OLV18] Oord, Aaron van den; Li, Yazhe; Vinyals, Oriol: Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748, 2018.

[PY10] Pan, Sinno Jialin; Yang, Qiang: A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10):1345–1359, 2010.

[Re17] Rebuffi, Sylvestre-Alvise; Kolesnikov, Alexander; Sperl, Georg; Lampert, Christoph H.: iCaRL: Incremental Classifier and Representation Learning. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. pp. 2001–2010, 2017.

[Re22] Reed, Colorado J.; Yue, Xiangyu; Nrusimha, Ani; Ebrahimi, Sayna; Vijaykumar, Vivek; Mao, Richard; Li, Bo; Zhang, Shanghang; Guillory, Devin; Metzger, Sean et al.: Self-supervised pretraining improves self-supervised pretraining. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 2584–2594, 2022.

[RJ18] Rahman, Shahina; Johnson, Valen E.: A Fast Algorithm for Clustering High Dimensional Feature Vectors. arXiv preprint arXiv:1811.00956, 2018.

[Ro87] Rousseeuw, Peter J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20:53–65, 1987.

[Sc23] Schliebitz, Andreas; Graf, Henri; Wamhof, Tobias; Tapken, Heiko; Gertzen, Andreas: KI-basiertes Computer-Vision-System zur Qualitäts- und Größenbestimmung von Kartoffeln. 43. GIL-Jahrestagung, Resiliente Agri-Food-Systeme, 2023.

[SFS16] Sun, Baochen; Feng, Jiashi; Saenko, Kate: Return of frustratingly easy domain adaptation. In: Proceedings of the AAAI conference on artificial intelligence. volume 30, 2016.

[SS18] Saravanan, R.; Sujatha, Pothula: A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification. In: 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS). pp. 945–949, 2018.

[To23] Torch Contributors: , Models and pre-trained weights. Online: `https://pytorch.org/vision/stable/models.html`, January 2023. Retrieved: 2023-08-11.

[Ts10] Tscherepanow, Marko: TopoART: A topology learning hierarchical ART network. In: Artificial Neural Networks–ICANN 2010: 20th International Conference, Thessaloniki, Greece, September 15-18, 2010, Proceedings, Part III 20. Springer, pp. 157–167, 2010.

[Wh09] Whitehill, Jacob; Wu, Ting-fan; Bergsma, Jacob; Movellan, Javier; Ruvolo, Paul: Whose vote should count more: Optimal integration of labels from labelers of unknown expertise. Advances in neural information processing systems, 22, 2009.

[ZW22] Zheng, Tianxiang; Wen, Zhijie: Online Convolutional Neural Network for Image Streams Classification. In: Proceedings of the 5th International Conference on Big Data Technologies. pp. 255–259, 2022.