BiCAE – A Bimodal Convolutional Autoencoder for Seed Purity Testing

Maksim Kukushkin [1], **Martin Bogdan**[1], **Thomas Schmid**[2,3,1]

[1]Leipzig University
[2]Martin Luther University Halle-Wittenberg
[3]Lancaster University in Leipzig
{kukushkin, bogdan}@informatik.uni-leipzig.de, thomas.schmid@medizin.uni-halle.de

Abstract

In the seed-producing industry, accurate assessment of harvested seeds for technical purity is a necessary, yet time-consuming and labor-intensive task. Automating this task holds immense potential for enhancing agricultural seed productivity, and using computer vision methods to classify seeds has already demonstrated promising results. Here, we propose a novel spectral-enhanced image anomaly detection approach to accurately discriminate Canola seeds (*Brassica napus* L.) from visually similar non-Canola seeds. Our bimodal approach exploits both RGB and data captured by a hyperspectral camera of the same sample. For efficient processing of this data, we suggest a novel bimodal convolutional autoencoder (BiCAE) architecture, which combines the strengths of high spatial resolution in RGB and high spectral resolution in hyperspectral data. We demonstrate that training our BiCAE model on a Canola dataset allows to learn a joint latent representation that effectively extracts spatio-spectral information from both RGB and hyperspectral data. Experiments show promising results in differentiating between Canola and non-Canola samples, in particular in detecting various types of non-Canola seeds in previously unseen test data. The obtained results highlight the model's ability to generalize beyond the training dataset, surpassing unimodal models that rely solely on a single modality.

Introduction

In the realm of agricultural seed production, ensuring seed quality is a major challenge. Avoiding the presence of unwanted seeds is not only in the interest of customers, but in many countries also subject to government regulations (Kuhlmann and Dey 2021; Batten, Plana Casado, and van Zeben 2021; Wattnem 2016). The European Union (EU), e.g., follows a relatively strict policy using ex ante quality control mechanisms (Batten, Plana Casado, and van Zeben 2021), such as restricting market access to seed lots that are properly certified (Winge 2015). In practice, this urges seed producing companies to analyze and classify harvested seeds on a regular basis. Achieving this on the legally required level, however, requires not only human analysts, but often necessitates extensive training for these experts.

It is therefore not surprising that researchers have investigated the use of deep learning techniques based on RGB data (see Section 2.1). However, such RGB-based approaches often overlook the fact that relying solely on color might not suffice for effectively distinguishing between various types of seeds. While RGB image analysis can offer valuable information about the visual properties of seeds, it cannot extract information about their chemical composition as it is limited to the visible spectrum. Consequently, this constraint frequently results in issues like metameric colors and the inability of models to accurately discriminate seed species.

As an alternative, approaches employing hyperspectral imaging instead of RGB data have been proposed in recent years, too (Feng et al. 2019; ElMasry et al. 2019). Hyperspectral imaging, however, poses its own challenges (Imani and Ghassemian 2020). It provides lower spatial resolution, resulting in a lack of capture of critical spatial details of seeds, such as their shape, surface patterns and irregularities. Furthermore, depending solely on hyperspectral data may cause challenges in model generalization. This is because the acquisition of hyperspectral images is susceptible to external factors such as lighting conditions, atmospheric influences, and sensor calibration, all of which can introduce variations that make models sensitive to these changes.

While utilizing either RGB or hyperspectral data provides the above mentioned challenges, researchers are in practice often limited to using only one of these modalities. For the work presented here, we are in the fortunate position to work with an industry partner that is able to provide both RGB and hyperspectral images recorded of the same seed. The data employed in this study originates from the agricultural production of Canola seed (*Brassica napus* L.). Canola seed is one of the most widely cultivated and essential oilseeds globally, with diverse applications in the food, chemical, automotive, and renewable energy industries (Arthey 2020). Our goal in this project is to differentiate Canola seeds from non-Canola seeds, where two different modalities are available. We propose a specialized anomaly detection system that utilizes both RGB and hyperspectral images for the given task and evaluate its effectiveness by analyzing its ability to identify samples from 12 commonly encountered weed species as non-Canola seed. The results of our evaluation of the BiCAE architecture hold great promise in improving efficiency and accuracy of technical seed purity analysis and the seed sorting processes.

Related Work

Computer Vision for Seed Analysis

To separate wanted from unwanted seed automatically, a number of technical solutions have been investigated, in particular using computer vision (Gong et al. 2015; Rahman and Cho 2016). A widely adopted approach is the classification of seed images based on labeled data, with numerous studies focusing on classification of different seed species such as rice (Qiu et al. 2018; Kiratiratanapruk et al. 2020), cottonseeds (Jamuna et al. 2010), sunflower (Luan et al. 2020; Barrio-Conde et al. 2023), tomato seeds (Škrubej et al. 2015), corn seeds (Taylor, Chiou, and Bond 2019; Ali et al. 2020), wheat (Agarwal, Bachan et al. 2023; Yasar 2023), plum kernels (Ropelewska et al. 2022), and maize (Bi et al. 2022). For Canola seeds, to the best of our knowledge only a study for bulk classification has been published so far (Qadri et al. 2021), but not for classification of individual seeds.

Most related studies employ machine learning (ML) approaches. In recent years, there has been a growing interest in utilizing deep learning techniques. Transfer learning, for instance, has been used to classify 14 common seed species (Gulzar et al. 2020; Hamid et al. 2022) and wheat varieties (Yasar 2023). Swin transformers have been used for classifying maize varieties (Bi et al. 2022), and AlexNet has been employed to classify sunflower seeds (Barrio-Conde et al. 2023). A recent development is to incorporate images from hyperspectral cameras instead of using traditional RGB images, which has yielded promising results so far (Feng et al. 2019; ElMasry et al. 2019; Fabiyi et al. 2020).

While the mentioned approaches are promising, they provide a major challenge for practical usage in the agricultural industry as they implement the task of detecting abnormal seeds as a classification task, requiring labeled data and being limited to pre-defined classes of seeds. From a practical point of view, approaches not relying on labeled abnormal samples – in other words: anomaly detection approaches – are clearly preferable as they implicitly are applicable to unknown abnormal samples. In previous work, we were already able to demonstrate that it is possible to implement this approach successfully in a unimodal fashion using either RGB or hyperspectral data (Kukushkin et al. 2023).

Anomaly Detection in Images. Detecting anomalies in images is of great significance and finds applications across various fields (Pimentel et al. 2014). Recently, there has been a growing interest in the development of unsupervised methods that do not rely on labeled data (Kiran, Thomas, and Parakkal 2018). Along with many machine learning models that perform well at spotting outliers (Perera, Oza, and Patel 2021), such as Isolation Forest (Liu, Ting, and Zhou 2008), One-Class SVM (Schölkopf et al. 1999), and Local Outlier Factor (Breunig et al. 2000), three main methods that use deep learning are also popular: autoencoders, generative adversarial networks, and transfer learning with convolutional neural networks (Pang et al. 2021).

Probably most popular and currently most widely used for anomaly detection in images, are autoencoders (AE) (Baldi 2012). For anomaly detection, AEs are trained with a set of normal images and then applied to analyze new, unseen images. Anomalies are identified when the AE struggles to accurately reconstruct an input image (Ribeiro, Lazzaretti, and Lopes 2018). Conventional AEs, however, may face challenges in reconstructing complex image features such as textures and patterns, leading to false positives or false negatives. To address this limitation, researchers have proposed various modifications to the standard AE architecture. For example, denoising autoencoders (Lu et al. 2017) and the inclusion of regularizers in the loss function (Vincent et al. 2008; Rifai et al. 2011; An and Cho 2015; Makhzani et al. 2015) have been suggested to enhance reconstruction accuracy and improve the performance of anomaly detection.

Neural Networks for Spectroscopic data. Spectroscopic data does typically not exhibit the same level of high-dimensional and highly non-linear relationships as image data (Swinehart 1962). Networks architectures designed for impedance (Schmid, Bogdan, and Günzel 2013), Raman (Liu et al. 2017), Near-infrared (NIR) (Cui and Fearn 2018), and XRD (Lee et al. 2020) spectroscopic data are therefore typically shallow, often limited to only a few layers. Attempts to adapt computer vision models, such as ResNet-50, to spectroscopic data have shown promise, outperforming traditional methods such as Random Forests (Lee et al. 2020). However, achieving perfect prediction accuracy remains a challenge, as demonstrated in a recent benchmark study (Schuetzke, Szymanski, and Reischl 2023).

Multimodal Autoencoders

In the domain of multimodal learning, several studies have explored the use of bimodal autoencoders, which incorporate multiple types of data. These models aim to learn a shared representation between different input modalities, enabling their utilization in various tasks. Ngiam et al. (Ngiam et al. 2011), for instance, proposed a bimodal deep autoencoder for speech recognition, leveraging both acoustic and visual data. Expanding on this concept, Sayed et al. (Sayed, ElDeeb, and Taie 2023) introduced a Variational Autoencoder (VAE) specifically designed for audiovisual data. Additionally, Nguyen et al. (Nguyen et al. 2021) employed two autoencoders with Long Short-Term Memory (LSTM) to extract audiovisual features for emotion recognition. Gong et al. (Gong et al. 2021) proposed a variational selective autoencoder for bimodal image data. For video classification tasks, researchers have utilized stacked autoencoders (Liu, Feng, and Zhou 2016), where individual stacked contractive autoencoders were created for each modality. The outputs of these encoders were then combined and fed into a multimodal stacked contractive autoencoder, allowing for the fusion of text, audio, and image modalities. Similarly, in video event detection (Jhuo and Lee 2014), fusion of multiple modalities was employed.

In summary, the integration of various types of autoencoders and multimodal fusion techniques has shown significant value in enhancing the capabilities of anomaly detection and multimodal learning systems. To the best of our knowledge, however, this study represents the first application of unsupervised learning techniques in the field of seed species anomaly detection, further expanding the range of methodologies employed in this domain.

Methodology

To enable the integration of RGB images as input, we standardized their dimensions by resizing or padding them to a uniform size of 192x192x3 pixels. Furthermore, to increase the computational efficiency within hyperspectral approaches, we extracted a region of interest (ROI) from the central area of the hyperspectral images. This selected ROI has dimensions of 12x12x300. This ROI was averaged along the first and second dimensions so that the final vector is 300x1, which allows focused data processing on the most informative region while significantly reducing computational complexity.

Unimodal Baselines

Considering the presence of two different types of data - hyperspectral data, characterized by high spectral resolution but low spatial resolution, and RGB data, characterized by low spectral resolution but high spatial resolution-we propose an effective approach for handling these data types. It's important to highlight that the selection of baselines was constrained by the high-dimensionality of our dataset, meaning that not all approaches (e.g. One-Class SVM, clustering methods) are suitable for our specific task.

Local Outlier Factor. We employ two specialized unimodal Local Outlier Factor (Breunig et al. 2000) models, tailored either for hyperspectral (HS-LOF) or RGB data (RGB-LOF). LOF identify instances in a dataset that deviate significantly from the local patterns of their neighboring points. It measures the degree of outlierness of a data point by comparing its density with that of its neighboring points.

Isolation Forest. The second category of baselines includes two unimodal tree-based Isolation Forest (Liu, Ting, and Zhou 2008) models. Analogous to the LOF baseline, we created individual models for each modality: HS-IForest and RGB-IForest. Isolation Forests are known for their rapid and efficient performance, making them particularly well-suited for extensive datasets like ours. For our experimentation, we employed the pyod (Zhao, Nasrullah, and Li 2019) implementation with a parameter of $n\_estimators = 500$.

Autoencoder. In the third category of baseline, we included two unimodal autoencoder-based models, namely RGB-AE and HS-AE (as shown in Fig. 1). In essence, each AE operates in two stages: Initially, it takes an input image r_i and compresses it into a lower-dimensional representation. In the subsequent stage, the AE aims to reconstruct the input image with the highest accuracy. To this end, it employs the Mean Squared Error (MSE) loss function for HS-AE (see Eq. 1) and Multi-Scale Structural Similarity Index (MS-SSIM) for RGB images (see Eq. 2), which both measures the difference between the reconstructed images r_i' and the input image r_i.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(r_i - r_i')^2 \tag{1}$$

$$MS\text{-}SSIM(r_i, r_i') = \prod_{i=1}^{N} SSIM_i(r_i, r_i')^{\beta_i} \tag{2}$$

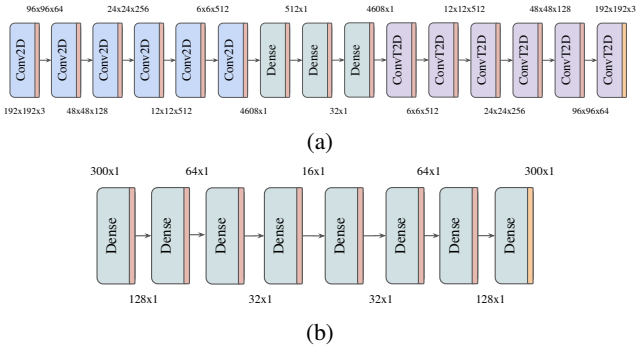

Figure 1: Architectures of RGB-AE (a) and HS-AE (b).

It is important to note that our HS-AE architecture differs from RGB-AE. HS-AE utilizes exclusively fully connected layers, effectively capturing the linear spectral features present in the spectra. Conversely, RGB-AE employs a ConvNet design with convolutional layers to extract spatial information, as depicted in Figure 1a.

We avoided using batch normalization between convolution and fully connected layers. This is because batch normalization tends to normalize both normal and anomalous data distributions, which could make the differences between the two classes less distinct. For both HS-AE and RGB-AE, we opted for the Gaussian Error Linear Unit (GELU) activation function, as introduced in (Hendrycks and Gimpel 2016). GELU provides a smoother activation function compared to ReLU, effectively capturing nonlinearity and improving performance in deeper networks.

Bimodal Convolutional Autoencoder (BiCAE)

In this section, we present our main contribution, which is a bimodal convolutional autoencoder called BiCAE. This model leverages the strengths of both hyperspectral (HS) and RGB data to enhance its performance (see Fig.2). Key features of the BiCAE are:

- **Combined Loss:** To effectively fuse information from both modalities and address the challenge of combining losses with significantly different ranges, we employ the geometric mean. The geometric mean has the ability to mitigate the influence of extreme values. Hence, the final loss for the BiCAE is formulated as follows:

$$Loss = \sqrt{MSE_{hs} \times MS\text{-}SSIM_{rgb}} \tag{3}$$

- **Fusion Module:** The fusion module is responsible for integrating the features from different modalities within the network. In line with the taxonomy presented by Baltrušaitis et al. (Baltrušaitis, Ahuja, and Morency 2018), we adopt the early-fusion approach. This involves concatenating the input features from different modalities.

By incorporating these features, the proposed bimodal convolutional autoencoder, namely BiCAE, which aims to achieve improved performance by effectively leveraging both HS and RGB data.

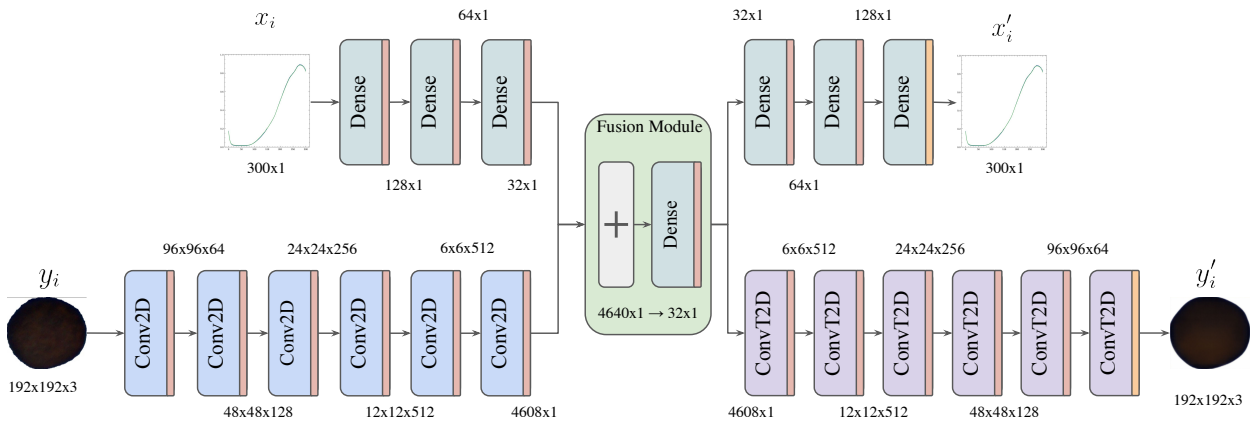

Figure 2: Overview our Bimodal Convolutional Autoencoder (BiCAE) architecture for detection of non-Canola seeds.

Experiments

In order to evaluate the reliability of our anomaly detection approach, we performed extensive tests using both Canola seeds and several types of non-Canola seeds.

Data

For evaluation of the proposed BiCAE architecture, we utilized a training dataset that included 4750 RGB images and their corresponding 4750 hyperspectral images of Canola. We set aside a separate dataset specifically for testing purposes, i.e., to assess the effectiveness of our approach. This test set comprised 4000 RGB images and 4000 hyperspectral images. Among these, 1000 images from each imaging modality belonged to the normal class, specifically representing the *Brassica napus* L. species. The remaining 600 images in the test set were evenly distributed among 12 distinct weed species: (i) *Anchusa arvensis* L., (ii) *Arctium lappa* L., (iii) *Alopecurus myosuroides* L., (iv) *Bistorta officinalis* L., (v) *Echinochloa crus-galli* L., (vi) *Fallopia convolvulus* L., (vii) *Galium aparine* L., (viii) *Geranium pratense* L., (ix) *Geranium dissectum* L., (x) *Galeopsis tetrahit* L., (xi) *Polygonum aviculare* L. and (xii) *Sinapis arvensis* L. (see Figure 3 and Table 1). The main objective of the evaluated anomaly detection task was to accurately identify these 1600 images from the anomalous class using an unsupervised learning.

The hyperspectral images were captured using the Resonon (USA) Pika L 100121-220 model. These images covered a range of 300 wavelengths within the visible and near-infrared (VNIR) segment of the electromagnetic spectrum, spanning from 380 nm to 1000 nm with a resolution of 5 nm (see Figure 4). The RGB images were obtained using the Sony (Japan) IMX477 model.

Training Setting

We trained both unimodal baseline models as well as our BiCAE architecture. Each model in our study was trained using the same configuration. This involved employing the Lion optimizer (Chen et al. 2023) with a learning rate of 0.0001 for 100 epochs, along with a learning rate schedule. To expand the training dataset, we employed data augmentation techniques, including vertical and horizontal flips, as well as rotations. Additionally, a batch size of 250 was employed during the training process.

The unimodal models, HS-AE and RGB-AE as well as HS-Branch and RGB-Branch, were trained using Mean Squared Error (MSE) loss, as depicted in Equation 1. However, for our proposed bimodal convolutional AE, BiCAE, we utilized a combined loss function (see Equation 3).

To assess and compare the performance of the two models, we employed several evaluation metrics. These metrics included: (i) the Area Under the ROC Curve (AUC), (ii) Average Precision (AP), (iii) Accuracy, (iv) Sensitivity, and (v) Specificity. To enhance the reliability of our findings, we train each model three times and evaluated it on an augmented test dataset (same augmentation as for train dataset) and then averaged the results to determine the final perfor-

Class	Species	Train	Test	$\sum$
Normal	*B. napus*	4750	1000	5750
Anomaly	*A. arvensis*	-	50	50
	A. lappa	-	50	50
	A. myosuroides	-	50	50
	B. officinalis	-	50	50
	E. crus-galli	-	50	50
	F. convolvulus	-	50	50
	G. aparine	-	50	50
	G. pratense	-	50	50
	G. dissectum	-	50	50
	G. tetrahit	-	50	50
	P. aviculare	-	50	50
	S. arvensis	-	50	50
$\sum$	13	4750	1600	6350

Table 1: Sample Allocation in Training and Test Sets per each (HS/RGB-) modality. The models under consideration are exclusively trained on *B.napus* seeds, with seeds from all other species categorized as the "anomaly" class.

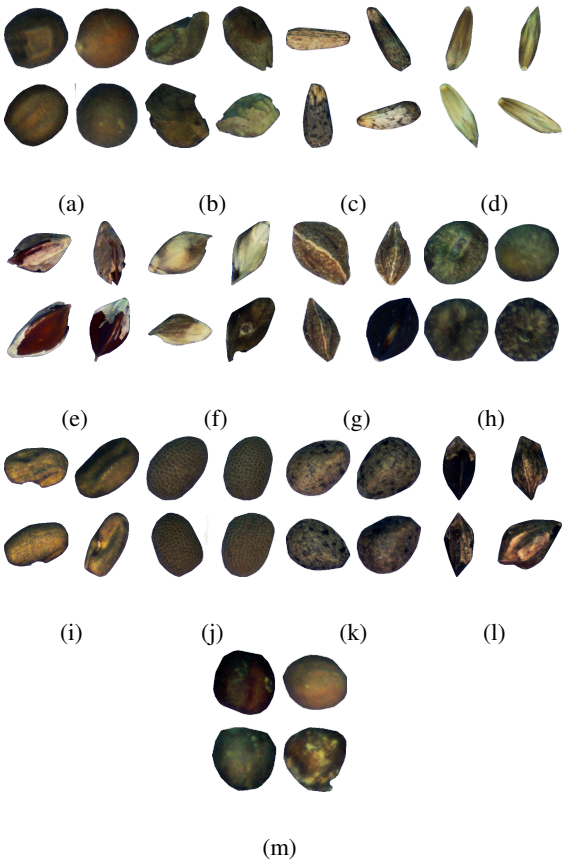

Figure 3: Example images from the test dataset. (a) Represents the normal class, showing images of *B. napus*, (b-m) Depict anomalous classes, displaying images of *A. arvensis* L., *A. lappa* L., *A. myosuroides* L., *B. officinalis* L., *E. crus-galli* L., *F. convolvulus* L., *G. aparine* L., *G. pratense* L., *G. dissectum* L., *G. tetrahit* L., *P. aviculare* L. and *S. arvensis* L. respectively.

mance. This should ensure that the models were evaluated not on identical data during each run.

The threshold for classification was determined using the ROC curve. We opted for this approach as the selection of the custom threshold can have a significant impact on metrics (ii) to (iv), leading to variations in their values.

Results

All trained anomaly detection models were evaluated on the test data using the metrics Accuracy (Acc.), Sensitivity (Sen.), Specifity (Spe.), Area Under Curve (AUC), and Average Precision (AP). Table 2 shows these metrics for both unimodal baselines as well as for our novel bimodal approach (BiCAE). The accuracy is compared visually in Figure 5, offering additional clarity.

Analyzing these results, it becomes evident that the BiCAE model outperformed all other methods across other metrics, exhibiting superior Accuracy (0.966), Sensitivity (0.985) and Specificity (0.955) in distinguishing between inliers and outliers.

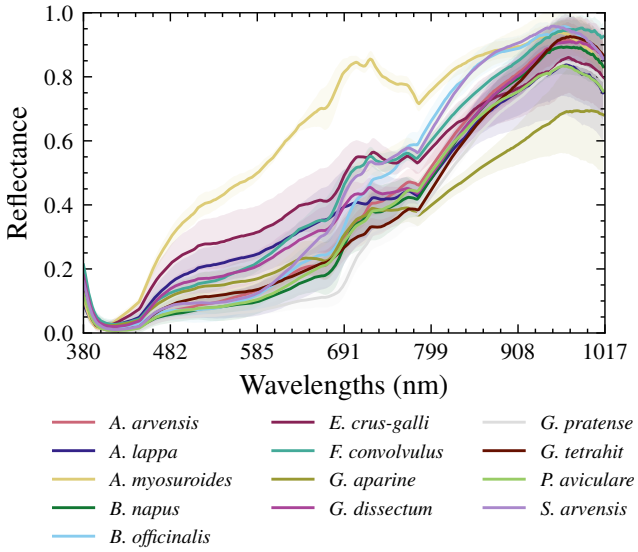

Figure 4: Spectra of the Region of Interest (ROI) for each seed species in the test dataset, illustrating the mean and standard deviation of a 12x12x300 dimensional ROI along first last dimension

Apart from the BiCAE, the models trained solely on hyperspectral data demonstrated better performance compared to their RGB counterparts. Notably, the lightweight HS-LOF showcasing relatively high Accuracy (0.965), Sensitivity (0.978) and Specificity (0.957). Conversely, its RGB-LOF counterpart displayed significantly lower values across Accuracy (0.871), Sensitivity (0.773) and Specificity (0.931), as depicted in the results table. The similar results are shown by HS-AE and RGB-AE. HS-AE outperformed RGB-AE and reached higher results in terms of Accuracy (0.959), Sensitivity (0.982) and Specificity (0.945).

Notably, RGB-IForest outperformed its hyperspectral

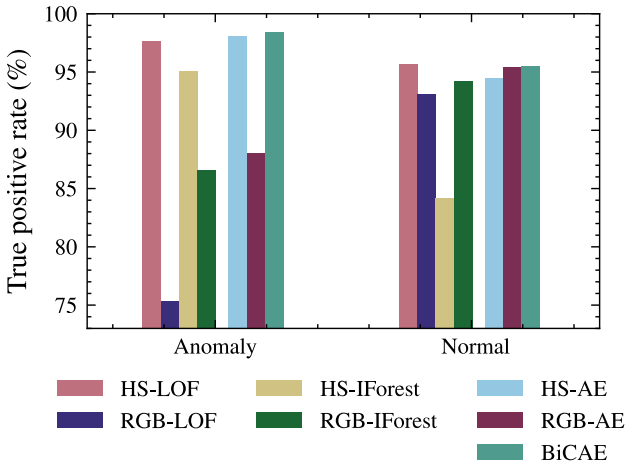

Figure 5: Average True Positive Rate and True Negative Rate for class (normal/anomaly)

Model	Accuracy	Sensitivity	Specificity	AUC	AP
HS-LOF	0.965±0.000	0.978±0.000	0.957±0.000	0.984±0.000	0.964 ±0.000
RGB-LOF	0.871±0.002	0.773±0.013	0.931±0.005	0.904±0.001	0.896±0.001
HS-IForest	0.883±0.009	0.952±0.005	0.842±0.016	0.953±0.001	0.905±0.001
RGB-IForest	0.918±0.001	0.877±0.006	0.942±0.006	0.960±0.001	0.949±0.001
HS-AE	0.959±0.002	0.982±0.004	0.945±0.001	0.989±0.000	0.980±0.000
RGB-AE	0.924±0.005	0.875±0.012	0.954±0.012	0.950±0.001	0.952±0.002
BiCAE	**0.966±0.004**	**0.985±0.007**	**0.955±0.010**	**0.995±0.000**	**0.992±0.001**

Table 2: Average test metrics across three runs for evaluated models

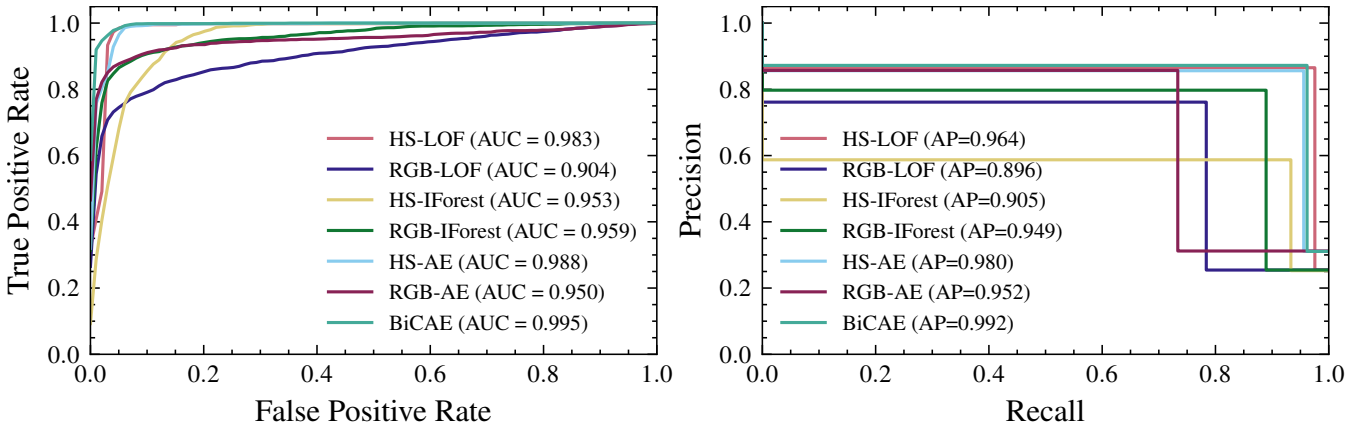

Figure 6: Performance comparison of considered models: (a) ROC curve and Area under the ROC curve, and (b) Precision-Recall curve and Average Precision.

analogue at most metrics, reaching higher Accuracy (0.918), lower Sensitivity (0.877), and higher Specificity (0.942).

Figure 6a and Figure 6b illustrate the performance of evaluated approaches. Among these, the BiCAE achieved an the highest Area Under the Curve score of 0.995, outperforming other models such as HS-AE (0.989), HS-LOF (0.984), HS-IForest (0.953), RGB-IForest (0.960) RGB-AE (0.950), and RGB-LOF (0.904). Moreover, the BiCAE demonstrated notably higher Average Precision values, securing a score of 0.992 across a range of potential decision thresholds.

Discussion

Model Comparison

Our BiCAE demonstrates remarkable reliability and effectiveness in accurately distinguishing between Canola and non-Canola samples, as evidenced by the results obtained from its training and testing phases. To assess and compare the performance of the seven models, we further analyzed two key metrics: AUC and AP. AUC is a widely used measure that evaluates a model's overall ability to differentiate between positive and negative samples across different decision thresholds. It is particularly valuable in balanced classification problems and ranking predictions. On the other hand, AP assesses the trade-off between precision and recall at various decision thresholds, making it useful in imbalanced classification scenarios.

Achieving a balance between precision and recall is crucial here, which makes AP particularly relevant. Therefore, when comparing these metrics, AP plays a more significant role in our anomaly detection task. When considering AUC and AP values of the evaluated models it becomes evident that our BiCAE model outperforms all unimodal models. The higher AUC score of BiCAE indicates its superior ability to discriminate between Canola and non-Canola samples across multiple decision thresholds. Further, the AP metric confirms these findings by highlighting the precision-recall trade-off achieved by the BiCAE in our experiment.

The significance of BiCAE's performance highlights the need to consider not only the visible information of RGB images, but also the chemical composition information that can be extracted using hyperspectral cameras. This is beneficial in cases where species have very similar shapes, color and texture, i.e. they appear identical to the human eye or trichromatic camera, similar texture or shape, so that using only the visible information will give insufficient results. Furthermore, our study demonstrated a case whereby the incorporation of hyperspectral information aided in distinguishing between such two species, namely *B. napus* and *S. arvensis*. To sum up, the superior performance of bimodal approach is

in line with the Ref. (ElMasry et al. 2019), which highlights the benefits of using hyperspectral data with RGB data in a broad range of applications.

Societal Impact

The development of autoencoder-based methods for detecting anomalous seed species holds promise for agricultural applications, in particular where current approaches are exceeded in terms of performance or reliability. At the same time, however, it is crucial to consider potential negative societal impacts that may arise from usage of such approaches in real-world scenarios. For the here described application, we would like to point in particular the following technological and practical aspects:

- **Bias and Discrimination:** AE models trained on biased or unrepresentative datasets can maintain biases and discrimination. This may result in misclassification of seed species, leading to technical impurities which might compromise legal thresholds and could disturb contractual agreements between seed suppliers and buyers or promoting inequitable agricultural practices, such as the excessive use of herbicides.
- **Overreliance on Automation:** While autoencoder-based systems can enhance efficiency, an excessive dependence on automation may undermine the role of agricultural experts and their valuable domain knowledge. It is important to include human expertise and manual inspection to ensure accurate identification of seed species. However, from our perspective, the integration of such systems is unlikely to result in the loss of many highly valuable seed analyst positions. This is primarily because manual seed classification is a extremely monotonous task that lacks ergonomic suitability.
- **Disruption of Traditional Agricultural Practices:** Implementing automated systems like autoencoder-based anomaly detection may disrupt traditional agricultural practices, affecting small-scale farmers or seed producers both socially and economically. Therefore, it is crucial to consider the potential impacts on local knowledge and livelihoods.

To mitigate these potential negative societal impacts, it is essential to adopt a responsible and inclusive approach. This involves several measures, including ensuring diverse and representative training datasets, addressing biases in data collection and model development, involving stakeholders in decision-making processes, and prioritizing transparency and accountability in the deployment of automated systems. Further, ongoing monitoring and evaluation of the socio-environmental consequences of such technology should be conducted to identify and mitigate any adverse effects.

Conclusion and Outlook

Our study has demonstrated that automating the differentiation between Canola and non-Canola seed can be improved using both RGB and hyperspectral images. We have successfully developed and evaluate a novel architecture that illustrate the effectiveness of autoencoders (AE) in detecting abnormal seeds using both RGB and hyperspectral data. We also investigated the conditions under which feature fusion is effective. To the best of our knowledge, this is the first study that applies anomaly detection in this context.

Notably, the remarkable performance of the BiCAE architecture underscores the significance of incorporating both RGB and hyperspectral data into a unified model. These findings indicate that our approach holds promise for improving and expediting not only Canola seed production but also agricultural seed production as a whole. The overall goal is to integrate our approach into the development of an AI-supported platform that combines advanced machine learning techniques with robotic sensors and actors.

Moving forward, our future research aims to advance anomaly detection techniques in agriculture. We plan to further investigate how two modalities can be combined within a bimodal AE, such as exploring methods for integrating losses from both modalities. Additionally, we intend to expand our dataset to include a wider range of seed species and evaluate the performance of our architecture on these new additions. Furthermore, we aim to identify the specific wavelengths that offer valuable information for distinguishing different types of plant seeds. This will simplify the anomaly detection process, improve efficiency, and potentially reduce the costs of agricultural seed production.

Acknowledgments

We express our gratitude to NPZ Innovation GmbH for generously providing the dataset. Moreover, we would like to thank Matthias Enders and Simon Goertz for discussions on various aspects of agricultural seed purity testing.

This project is supported by funds from the German Federal Ministry of Food and Agriculture (BMEL), based on a decision of the Parliament of the Federal Republic of Germany. The German Federal Office for Agriculture and Food (BLE) provides coordinating support for artificial intelligence (AI) in agriculture as the funding organization, grant number 28DK116C20.

References

Agarwal, D.; Bachan, P.; et al. 2023. Machine learning approach for the classification of wheat grains. *Smart Agricultural Technology*, 3: 100136.

Ali, A.; Qadri, S.; Mashwani, W. K.; Brahim Belhaouari, S.; Naeem, S.; Rafique, S.; Jamal, F.; Chesneau, C.; and Anam, S. 2020. Machine learning approach for the classification of corn seed using hybrid features. *International Journal of Food Properties*, 23(1): 1110–1124.

An, J.; and Cho, S. 2015. Variational autoencoder based anomaly detection using reconstruction probability. *Special lecture on IE*, 2(1): 1–18.

Arthey, T. 2020. Challenges and perspectives in global rapeseed production. *Available on line at: http://www. agribenchmark. org/cash-crop/publicationsand-projects0/reports/challenges-and-perspectives-in-global-rapeseedproduction. html*.

Baldi, P. 2012. Autoencoders, unsupervised learning, and deep architectures. In *Proceedings of ICML workshop on unsupervised and transfer learning*, 37–49. JMLR Workshop and Conference Proceedings.

Baltrušaitis, T.; Ahuja, C.; and Morency, L.-P. 2018. Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis and machine intelligence*, 41(2): 423–443.

Barrio-Conde, M.; Zanella, M. A.; Aguiar-Perez, J. M.; Ruiz-Gonzalez, R.; and Gomez-Gil, J. 2023. A Deep Learning Image System for Classifying High Oleic Sunflower Seed Varieties. *Sensors*, 23(5): 2471.

Batten, L.; Plana Casado, M. J.; and van Zeben, J. 2021. Decoding seed quality: a comparative analysis of seed marketing law in the EU and the United States. *Agronomy*, 11(10): 2038.

Bi, C.; Hu, N.; Zou, Y.; Zhang, S.; Xu, S.; and Yu, H. 2022. Development of deep learning methodology for maize seed variety recognition based on improved swin transformer. *Agronomy*, 12(8): 1843.

Breunig, M. M.; Kriegel, H.-P.; Ng, R. T.; and Sander, J. 2000. LOF. *ACM SIGMOD Record*, 29(2): 93–104.

Chen, X.; Liang, C.; Huang, D.; Real, E.; Wang, K.; Liu, Y.; Pham, H.; Dong, X.; Luong, T.; Hsieh, C.-J.; Lu, Y.; and Le, Q. V. 2023. Symbolic Discovery of Optimization Algorithms.

Cui, C.; and Fearn, T. 2018. Modern practical convolutional neural networks for multivariate regression: Applications to NIR calibration. *Chemometrics and Intelligent Laboratory Systems*, 182: 9–20.

ElMasry, G.; Mandour, N.; Al-Rejaie, S.; Belin, E.; and Rousseau, D. 2019. Recent applications of multispectral imaging in seed phenotyping and quality monitoring—An overview. *Sensors*, 19(5): 1090.

Fabiyi, S. D.; Vu, H.; Tachtatzis, C.; Murray, P.; Harle, D.; Dao, T. K.; Andonovic, I.; Ren, J.; and Marshall, S. 2020. Varietal Classification of Rice Seeds Using RGB and Hyperspectral Images. *IEEE Access*, 8: 22493–22505.

Feng, L.; Zhu, S.; Liu, F.; He, Y.; Bao, Y.; and Zhang, C. 2019. Hyperspectral imaging for seed quality and safety inspection: A review. *Plant methods*, 15(1): 1–25.

Gong, Y.; Hajimirsadeghi, H.; He, J.; Durand, T.; and Mori, G. 2021. Variational selective autoencoder: Learning from partially-observed heterogeneous data. In *International Conference on Artificial Intelligence and Statistics*, 2377–2385. PMLR.

Gong, Z.; Cheng, F.; Liu, Z.; Yang, X.; Zhai, B.; and You, Z. 2015. Recent developments of seeds quality inspection and grading based on machine vision. In *2015 ASABE Annual International Meeting*, 1. American Society of Agricultural and Biological Engineers.

Gulzar, Y.; Hamid, Y.; Soomro, A. B.; Alwan, A. A.; and Journaux, L. 2020. A convolution neural network-based seed classification system. *Symmetry*, 12(12): 2018.

Hamid, Y.; Wani, S.; Soomro, A. B.; Alwan, A. A.; and Gulzar, Y. 2022. Smart seed classification system based on MobileNetV2 architecture. In *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, 217–222. IEEE.

Hendrycks, D.; and Gimpel, K. 2016. Gaussian Error Linear Units (GELUs).

Imani, M.; and Ghassemian, H. 2020. An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges. *Information fusion*, 59: 59–83.

Jamuna, K.; Karpagavalli, S.; Vijaya, M.; Revathi, P.; Gokilavani, S.; and Madhiya, E. 2010. Classification of seed cotton yield based on the growth stages of cotton crop using machine learning techniques. In *2010 International Conference on Advances in Computer Engineering*, 312–315. IEEE.

Jhuo, I.-H.; and Lee, D. 2014. Video event detection via multi-modality deep learning. In *2014 22nd International Conference on Pattern Recognition*, 666–671. IEEE.

Kiran, B. R.; Thomas, D. M.; and Parakkal, R. 2018. An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos. *Journal of Imaging*, 4(2): 36.

Kiratiratanapruk, K.; Temniranrat, P.; Sinthupinyo, W.; Prempree, P.; Chaitavon, K.; Porntheeraphat, S.; and Prasertsak, A. 2020. Development of paddy rice seed classification process using machine learning techniques for automatic grading machine. *Journal of Sensors*, 2020.

Kuhlmann, K.; and Dey, B. 2021. Using regulatory flexibility to address market informality in seed systems: A global study. *Agronomy*, 11(2): 377.

Kukushkin, M.; Enders, M.; Kaschuba, R.; Bogdan, M.; and Schmid, T. 2023. Canola seed or not? Autoencoder-based Anomaly Detection in Agricultural Seed Production. In *INFORMATIK 2023 - Designing Futures: Zukünfte gestalten*, 1645–1652. Bonn: Gesellschaft für Informatik e.V. ISBN 978-3-88579-731-9.

Lee, J.-W.; Park, W. B.; Lee, J. H.; Singh, S. P.; and Sohn, K.-S. 2020. A deep-learning technique for phase identification in multiphase inorganic compounds using synthetic XRD powder patterns. *Nature Communications*, 11(1).

Liu, F. T.; Ting, K. M.; and Zhou, Z.-H. 2008. Isolation Forest. In *2008 Eighth IEEE International Conference on Data Mining*. IEEE.

Liu, J.; Osadchy, M.; Ashton, L.; Foster, M.; Solomon, C. J.; and Gibson, S. J. 2017. Deep convolutional neural networks for Raman spectrum recognition: a unified solution. *The Analyst*, 142(21): 4067–4074.

Liu, Y.; Feng, X.; and Zhou, Z. 2016. Multimodal video classification with stacked contractive autoencoders. *Signal Processing*, 120: 761–766.

Lu, W.; Cheng, Y.; Xiao, C.; Chang, S.; Huang, S.; Liang, B.; and Huang, T. 2017. Unsupervised sequential outlier detection with deep architectures. *IEEE transactions on image processing*, 26(9): 4321–4330.

Luan, Z.; Li, C.; Ding, S.; Wei, M.; and Yang, Y. 2020. Sunflower seed sorting based on convolutional neural network. In *Eleventh International Conference on Graphics and Image Processing (ICGIP 2019)*, volume 11373, 428–434. SPIE.

Makhzani, A.; Shlens, J.; Jaitly, N.; Goodfellow, I.; and Frey, B. 2015. Adversarial autoencoders. *arXiv preprint arXiv:1511.05644*.

Ngiam, J.; Khosla, A.; Kim, M.; Nam, J.; Lee, H.; and Ng, A. Y. 2011. Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, 689–696.

Nguyen, D.; Nguyen, D. T.; Zeng, R.; Nguyen, T. T.; Tran, S. N.; Nguyen, T.; Sridharan, S.; and Fookes, C. 2021. Deep auto-encoders with sequential learning for multimodal dimensional emotion recognition. *IEEE Transactions on Multimedia*, 24: 1313–1324.

Pang, G.; Shen, C.; Cao, L.; and Hengel, A. V. D. 2021. Deep learning for anomaly detection: A review. *ACM computing surveys (CSUR)*, 54(2): 1–38.

Perera, P.; Oza, P.; and Patel, V. M. 2021. One-Class Classification: A Survey.

Pimentel, M. A.; Clifton, D. A.; Clifton, L.; and Tarassenko, L. 2014. A review of novelty detection. *Signal processing*, 99: 215–249.

Qadri, S.; Furqan Qadri, S.; Razzaq, A.; Ul Rehman, M.; Ahmad, N.; Nawaz, S. A.; Saher, N.; Akhtar, N.; and Khan, D. M. 2021. Classification of canola seed varieties based on multi-feature analysis using computer vision approach. *International Journal of Food Properties*, 24(1): 493–504.

Qiu, Z.; Chen, J.; Zhao, Y.; Zhu, S.; He, Y.; and Zhang, C. 2018. Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network. *Applied Sciences*, 8(2): 212.

Rahman, A.; and Cho, B.-K. 2016. Assessment of seed quality using non-destructive measurement techniques: a review. *Seed Science Research*, 26(4): 285–305.

Ribeiro, M.; Lazzaretti, A. E.; and Lopes, H. S. 2018. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recognition Letters*, 105: 13–22.

Rifai, S.; Vincent, P.; Muller, X.; Glorot, X.; and Bengio, Y. 2011. Contractive auto-encoders: Explicit invariance during feature extraction. In *Proceedings of the 28th international conference on international conference on machine learning*, 833–840.

Ropelewska, E.; Cai, X.; Zhang, Z.; Sabanci, K.; and Aslan, M. F. 2022. Benchmarking machine learning approaches to evaluate the cultivar differentiation of plum (prunus domestica L.) kernels. *Agriculture*, 12(2): 285.

Sayed, H. M.; ElDeeb, H. E.; and Taie, S. A. 2023. Bimodal variational autoencoder for audiovisual speech recognition. *Machine Learning*, 112(4): 1201–1226.

Schmid, T.; Bogdan, M.; and Günzel, D. 2013. Discerning apical and basolateral properties of HT-29/B6 and IPEC-J2 cell layers by impedance spectroscopy, mathematical modeling and machine learning. *PLOS ONE*, 8(7).

Schölkopf, B.; Williamson, R. C.; Smola, A.; Shawe-Taylor, J.; and Platt, J. 1999. Support Vector Method for Novelty Detection. In Solla, S.; Leen, T.; and Müller, K., eds., *Advances in Neural Information Processing Systems*, volume 12. MIT Press.

Schuetzke, J.; Szymanski, N. J.; and Reischl, M. 2023. Validating neural networks for spectroscopic classification on a universal synthetic dataset. *npj Computational Materials*, 9(1).

Škrubej, U.; Rozman, Č.; Stajnko, D.; et al. 2015. Assessment of germination rate of the tomato seeds using image processing and machine learning. *European Journal of Horticultural Science*, 80(2): 68–75.

Swinehart, D. F. 1962. The Beer-Lambert Law. *Journal of Chemical Education*, 39(7): 333.

Taylor, J.; Chiou, C.-P.; and Bond, L. J. 2019. A methodology for sorting haploid and diploid corn seed using terahertz time domain spectroscopy and machine learning. In *AIP Conference Proceedings*. Author(s).

Vincent, P.; Larochelle, H.; Bengio, Y.; and Manzagol, P.-A. 2008. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, 1096–1103.

Wattnem, T. 2016. Seed laws, certification and standardization: outlawing informal seed systems in the Global South. *The Journal of Peasant Studies*, 43(4): 850–867.

Winge, T. 2015. Seed legislation in Europe and crop genetic diversity. *Sustainable Agriculture Reviews: Volume 15*, 1–64.

Yasar, A. 2023. Benchmarking analysis of CNN models for bread wheat varieties. *European Food Research and Technology*, 249(3): 749–758.

Zhao, Y.; Nasrullah, Z.; and Li, Z. 2019. PyOD: A Python Toolbox for Scalable Outlier Detection. *Journal of Machine Learning Research*, 20(96): 1–7.