Highlights

Nondestructive in-ovo sexing of Hy-Line Sonia eggs by EggFormer using hyperspectral imaging

Chengming Ji, Ke Song, Zixin Chen, Shanyong Wang, Huanliang Xu, Kang Tu, Leiqing Pan, Junxian Huang

- Hyperspectral images of Hy-Line Sonia eggs were collected on even days from day 0 to 14 for gender identification.
- Feature bands were selected by RF, PCA, SPA, CARS and then the recombined images were processed by ViT-Base16.
- EggFormer demonstrates superior accuracy of 95.4%, with f1 score of 0.958 amd Kappa of 0.908 on day 10.
- By interpreting EggFormer, images with full bands were reduced to 22 bands retaining the same results, less than the 25 bands extracted by CARS.

Nondestructive in-ovo sexing of Hy-Line Sonia eggs by EggFormer using hyperspectral imaging

Chengming Ji[a], Ke Song[b], Zixin Chen[a], Shanyong Wang[c], Huanliang Xu[a], Kang Tu[b], Leiqing Pan[b,*], Junxian Huang[a,*]

[a] *College of Artificial Intelligence, Nanjing Agricultural University, Nanjing, 210095, Jiangsu, China*
[b] *College of Food Science and Technology, Nanjing Agricultural University, Nanjing, 210095, Jiangsu, China*
[c] *School of Information Science and Technology, ShanghaiTech University, 201210, Shanghai, China*

Abstract

Early identification of egg gender during incubation is crucial for animal welfare and commercial poultry production, as nowadays day-old male chicks are often culled due to low economic value. Hyperspectral imaging (HSI) recognition presents a swift, non-destructive, and cost-effective solution for in-ovo sexing compared to traditional methods such as Polymerase Chain Reaction (PCR), Volatile Organic Compounds (VOC), and Raman spectroscopy. In this study, we collected spectral images of Hy-Line Sonia chicken eggs even-numbered day from day 0 to 14, with a focus on day 10 for detailed analysis. We introduced the EggFormer model, incorporating channel attention and transformer self-attention mechanisms. To assess model performance, significant wavelengths were extracted by machine learning algorithms, including Random Forest(RF), Principal Component Analysis(PCA), Successive Projections Algorithm (SPA), and Competitive Adaptive Reweighted Sampling Algorithm (CARS). The channel images of these significant wavelengths were then employed with ViT-Base(Vision Transformer) for prediction and comparison. The EggFormer model demonstrated superior results, with accuracy of 95.4% , recall of 98.6%, f1 score of 0.958 and Kappa of 0.908. Besides, by interpreting channel attention block, 22 wavelengths were selected maintaining best results and 4 bands with accuracy of 94.6%. This outperformance positions it as a promisingly efficient and economical solution for industrial applications. The code of this work is available at https://github.com/quietbamboo/EggFormer for reproducibility.

Keywords: hyperspectral imaging, in-ovo sexing, deep learning, interpretable

1. Introduction

Worldwide, approximately 7 billion unwanted day-old male layer chicks are culled hours after they hatch every year(Krautwald-Junghanns et al., 2018), and 330 million in the European Union alone(Jia et al., 2023), sparking the ethical and animal welfare concerns(He et al., 2019). Since January 1, 2022, Germany and France have jointly become the first countries to ban the systematic killing of male chicks, and more and more countries are responding with such bans(Di Concetto et al., 2023). Some believe that chicken embryos potentially start to perceive pain after day 7 of the total 21-day incubation period before hatching(Rosenbruch, 1997), but encephalogram signals are not visible until after day 12, indicating that the pain perception may have occured after that period(Corion et al., 2023, Mellor and Diesch, 2007, Corion et al., 2022). In general, terminating chicken embryos before day 14 are generally believed to reduce the pain perception. Besides the animal welfare concerns, gender identification of eggs as early as possible during incubation can also substantially lower the per female layer chick cost and increase the overall hatching production efficiency. Moreover, male eggs identified during incubation can be used for other purposes, for example, male egges at day 10 of incubation may be used for vaccine production, while continuing with hatching not only incurs additional cost in the subsequent hatching process, but also results in a much lower economic value per egg, because a day 1 male chick is often immediately killed and used as a cheap protein source, such as in pet food.

Raman spectroscopy has been used for in-ovo sexing, achieving a 90% identification accuracy at day 3.5 of incubation by analyzing the spectra of blood in the extraembryonic vessels by opening a window in the shell(Galli et al., 2016). However, such invasive identification methods are less preferrable than noninvasive ones, due to increased contamination risks and possiblly lower hatchability rate. Moverover, the process of opening and sealing the eggshells would increase the cost of identification, making it unsuitable for large-scale applications in large hatcheries.

Volatile organic compounds(VOC)(Jia et al., 2023) in the gases emitted through eggshells are studied with promising in-ovo sexing results(Corion et al., 2023, Hu et al., 2022), however, collecting gas samples is time consuming and may potentially hurt the embryos since faster

* Corresponding author. E-mail address: jim@njau.edu.cn
First Author and Second Author contribute equally to this work.

gas extraction may result in lower oxygen levels inside the eggshell. Although recently techniques, such as chemical ionization have been proposed for rapid detection of gas components, the accuracy and repeatability of the model are still poorly understood and uncertain.

Compared with the above methods, spectral-based technology has the advantage of higher accuracy, higher throughput, and lower cost(Jia et al., 2023). Therefore, spectral-based noninvasive in-ovo sexing methods has become a trending research topic in poultry breeding. Hyperspectral imaging (HSI) technology can obtain spectral and image information at a series of wavelengths. Spectral information can reflect the composition of substances, such as protein, fat, moisture, etc., and image information can reflect external quality and surface defects. Due to the changes of eggshell and internal components caused by physiological metabolism during storage or incubation, hyperspectral images will also demonstrate differences, which serve as a strong basis for early embryo gender identification.

HSI based method is widely believed to be an ideal solution for in-ovo sexing(Göhler et al., 2017, Pan et al., 2016, Rahman et al., 2022, Corion et al., 2022). Typically, it covers three wavelength regions: near-ultraviolet(300-380nm), visible(380-780nm) and near-infrared(780-1500nm). It has been used to identify the unfertilized duck eggs(Dong et al., 2019) and chickena eggs(Ghaderi et al., 2024) before hatching, with over 90% prediction accuracy with visible/near infrared(VIS/NIR) transmittance spectroscopy.

While measuring full spectrum of wavelength may produce high accuracy, the increased measurement time and building cost of the full spectrum camera poses challenges. Hence it is important to extract feature bands among all the measured wavelengths for faster implementation and lower building cost with similar accuracy. Machine learning(ML) algorithms are proposed to select such feature bands, including Random Forest(RF)(Toksoz et al., 2021), Principal Component Analysis(PCA)(Corion et al., 2022, Galli et al., 2017), Successive Projections Algorithm(SPA)(Jia et al., 2023) and Competitive Adaptive Reweighted Sampling Algorithm(CARS)(Jia et al., 2023) Specifically, Deel Learning(DL) algorithms have been used in RGB images for gender identification(Horkaew et al., 2024, Jia et al., 2023), however few works have been designed for hyperspectral images.

Vision Transformer(ViT)(Dosovitskiy et al., 2020) based on Transformer(Vaswani et al., 2017), is a popular method for computer vision tasks for the usual 3-channel RGB images. In this study, we proposed a new model EggFormer based on ViT to identify the gender of eggs by HSI during hatching. Compared to conventional algorithms such as ML algorithms including RF, PCA, SPA, CARS, and DL algorithms including ViT-Base/16, EggFormer achieves state-of-the-art performance with the overall accuracy of 94.6% and precision of 94.8% on day 10. This work sheds light on solving the problem of in-ovo-sexing using the Transformer framework. We believe that our study can greatly boost the development of the new generation of high-throughput automated machines for in-ovo sexing, which can both benefit animal welfare and increase the efficiency for the poultry industry.

2. Materials and methods

2.1. Data collection and prerocessing

2.1.1. Materials and devices

Materials: The experimental materials were processed and collected in Nanjing Agricultural University, Jiangsu, China. Hy-Line Sonia eggs (white shell) were purchased from Jiushan Agriculture and Animal Husbandry Crop, Hubei, China.

Devices: The microcomputer automatic incubator (WSGD-6, Wansheng Incubation Equipment Crop, Nanjing, China); The Vision-near-infrared hyperspectral imaging system (Isuzu Optics Corp, Taiwan, China).

2.1.2. Sample preprocessing

A total of 180 Hy-Line Sonia eggs were utilized in all experiments, selecetd by similar color and size(60 ± 5 g) with no cracks. The eggs were then cleaned, sterilized with 75% alcohol, kept dry, and automatically turned every 2 hours inside incubator with environment maintaing at 37.8 °C and 60% relative humidity. Hyperspectral images of the pointed end of the eggs were collected even days from day 0 to 14. During this period, the eggs were immediately returned to the incubator after images collection to avoid any impact on the survival rate. On the 18th day of incubation, the eggs were transferred to a chamber until the chicks emerged from the shells on day 21 for furture gender identification.

2.1.3. Day-old chick gender identification

For Hy-Line Sonia, the gender of day-old chicks can be determined based on the length of feathers. For example, the covert feathers of female chicks are shorter than the primary feathers. In contrast, male chicks exhibit longer covert feathers compared to primaries, or in the same length. Finally, 70 male samples and 61 female samples were identified among the 131 hatched samples by the feathers sexing identification method.

2.1.4. HSI collection and correction

The hyperspectral transmission images of eggs were collected by line scanning method. The dull end of the egg was placed facing up, the light source was located directly below and the spectral camera was located above 30 cm, so the light transmit the egg and finally collected by camera. The physical picture of the hyperspectral acquisition system is shown in Figure 1.

The hyperspectral imaging system was firstly preheated for 30 minutes, and then after image collection lasting about 30 s, each egg was immediately put back into the incubator to keep intact. Meanwhile, to ensure the quality of the acquired hyperspectral images, it is necessary to optimize the exposure time, platform moving speed, light intensity and other parameters before the test. The main parameters finnally used of the imaging system are shown in Table 1. Due to the unstable intensity distribution of the light source of CCD camera, the hyperspctral image correction was carried out with black and white correction method(Cao et al., 2022).

Table 1: Main parameter settings of the imaging system

Parameters	Value (unit)
Image Resolution	440*804(pixel)
Acquisition Speed	1.5 (mm/s)
Spectral Resolutio	2.8 (nm)
Light Source Color	Yellow
Light Source Intensity	90 (W)
Color Temperature Range	1500 - 3500(K)
Exposure Time	72 (ms)
Wavelength Scope	382.67 - 1010.65(nm)

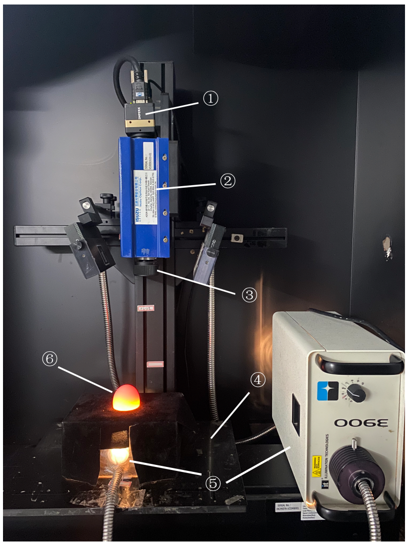


Figure 1: Physical diagram of hyperspectral acquisition system. ①Industrial camera; ②Imaging spectrometer; ③Lens; ④Moving platform; ⑤Halogen light source; ⑥Egg sample.

2.2. Region of interest slicing

The wavelength scope of collected spectral images ranges from 382.67nm to 1010.65nm, and the resolution after correction is 440(bands) * 804 pixels(width) * 377 pixels(height). However, not whole image contains egg, so it is necessary to slice the region of interest(ROI) from background. Specifically, the extracting flow of ROI was described in Figure 2, and it shows that the original image contains lots of noise information around the egg. Firstly, RGB image was merged by bands 638.82nm(R), 548.83nm(G) and 459.64nm(B), then greyscale image was processed by OpenCV library. The scope of ROI can be obtained directly by HoughCircle detection method, thus ROI spectral images can be sliced. The ROI radius was expanded by 5 pixels since the embryos were observed on the edge and the egg tends not to be a standard circle. Although little background noise was introduced, the potential information of egg border was more important.

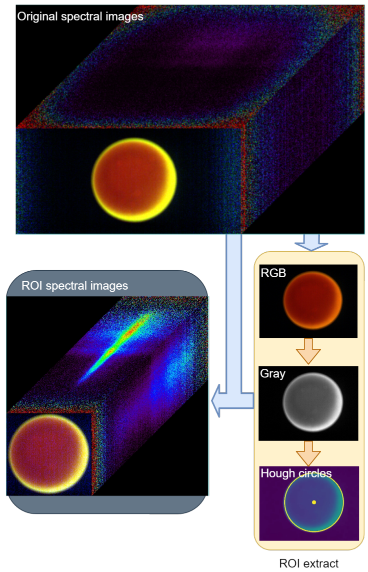


Figure 2: Extracting flow of region of interest(ROI).

2.3. Significant wavelengths selecting

2.3.1. Dimensional reduction methods

For the original spectral images contains 440 wavelengths, which caused the difficulty of training models as input, so the dimensional reduction is important. Random Forest(RF)(Cao et al., 2022, Belgiu and Drăguţ, 2016) and Principal Component Analysis(PCA)(Shahin and Symons, 2011) are typical methods to preselect the most relevant and significant wavelengths from hyperspectral images. Detailly, Random Forest is an ensemble classifier consisting of multiple decision trees, each of which is an estimator, then the result is computed by voting. Hence, RF analyses the importance of input variables, and then selecting the significant wavelengths (Strobl et al., 2008).

PCA is also a well-know method for feature extraction (Cao et al., 2003, Hasan and Abdulazeez, 2021), which transforms high-dimensional inputs into low-dimensional outputs. In fact, the principal components computed by PCA is consisted by input features, and thus the contribution value of input features (eigenvectors) can be referenced to select significant wavelength (Shahin and Symons, 2011).

Besides, Successive projections algorithm(SPA) is an algorithm which calculates correlation by projecting the vector representing the wavelength onto other wavelengths and comparing the projection magnitudes(Soares et al., 2013, Sun et al., 2019). It selects wavelength variable combinations with minimal redundant information and minimal collinearity.

Competitive adaptive reweighted sampling algorithm (CARS) is another widely used method to remove redundant information from HSI(Li et al., 2009, Liu et al., 2020). CARS is base on adaptive reweighted sampling (ARS) technique, which utilizes Partial Least Squares (PLS) modeling to compare regression coefficient magnitudes for wavelength selection. Through iterative cycles, the algorithm identifies the wavelength combination with the minimum RMSECV score, achieving optimal selection.

2.3.2. Average representative spectrums

Traditionally, for each hyperspectral image, a representative spectrum can be computed by the average of all pixels in each channel image, so each egg sample can be described as a 440-dimensional vector. The representative average spectrums of all eggs classified by female and male were shown in Figure 3. The relationship between wavelength and transmission value on day 0, 2, 4, 6 and day 8, 10, 12, 14 were descirbed in Figure (a) and (b), respectively. All curves contains two peaks around 700 nm and 800 nm, and the wavelengths at the peaks increase as time goes on. Specially, the peak value of female curve is higher than male curve on the same day. This phenomenon illuminates the female eggs are generally weaker in absorption than male eggs, so the value of transimitted light is higher, and the changes during incubation can be captured by hyperspectral imges.

2.3.3. Spectrums preprocessing methods

In order to eliminate the noise in the spectral images, following methods were used to smooth the spectrums or extract the features of spectrums, such as: savitzjy golay(SG)(Ai et al., 2022), multiplicative scatter correction(MSC) (Ma et al., 2016), standard normalized variate(SNV)(Guo et al., 1999), 1st and 2nd derivatives (Femenias et al., 2021). The smoothed results on day 10 were showed as example in Figure 4. However, little difference was found between the results of original and SG(window length=5, polyorder=3) smoothed spectrums, which indicating that great noise have been reduced after black-and-white correction and enough information contained in original spectrums, rendering it ideal for subsequent analysis.

2.4. Egg-SpectorFormer for in-ovo sexing

The conventional average spectrum method may inevitably lose some hidden information because the image of each channel is calculated as just a mean value. It is a great challenge for both machine learning(ML) algorithms and deep learning(DL) algorithms to use full bands of iamges as input for huge parameters and calculations. Thus, We proposed the EggFormer base on Vit, with Squeeze and Extract(SE) Layer and Depthwise Separable Convolution(DWConv), making it be capable of extracting more potential information from full bands input, and the model structure was shown in Figure 5.

2.4.1. SE Layer

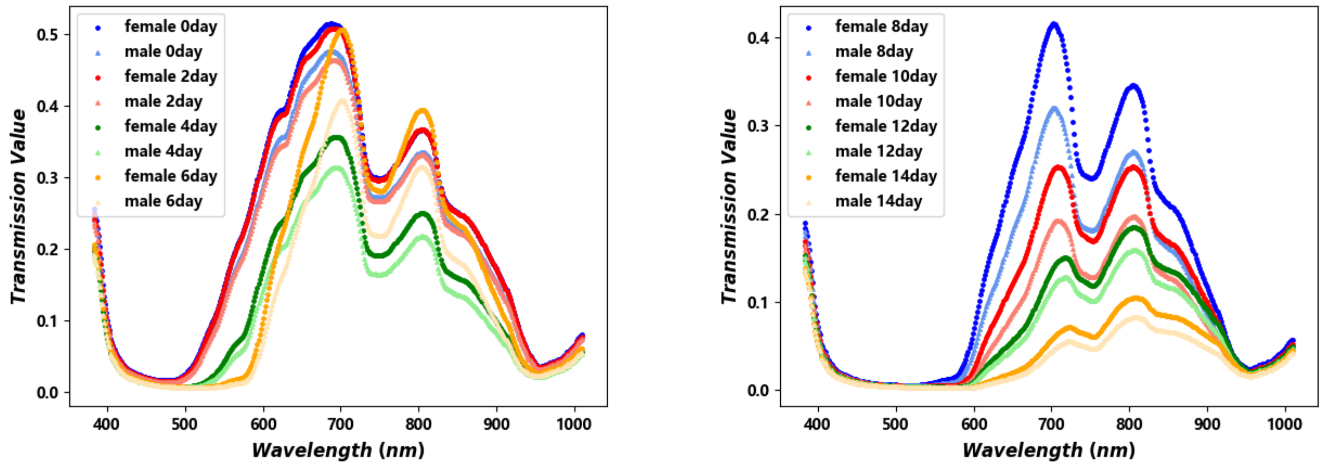
The SE layer(Hu et al., 2018) is a channel attention mechanism, which assigns larger weights to important channels and then linearly combines these channels based on the weights. Specifically, by using AvgPool2d, each channel of the spectral image is downsampled to a mean value, to considering more informations, the d1 and d2 value were also calculated in this work. Then the 3 440-dimensional vectors passed through two fully connected layers (Linear) to obtain weights for each channel. After that, the values were added by position and resized between 0 and 1 by sigmoid, and finally, the weights are multiplied with the corresponding channels images, resulting in the incorporation of channel importance features in the output data. During the training process, the parameters of 2 Linears are updated along with other model parameters until the correct channel importance weights are extracted.

2.4.2. Depthwise Separable Convolution

After SE layer, the output data keeps the same dimension with input, which contains some redundant information in some unimportant channels. To reduce the image channels, DWConv(Fran et al., 2017) is employed. It combines depthwise(DW) and pointwise(PW) components to extract features, resulting in lower parameter count and computational cost compared to conventional convolution. In the DWConv block, the image is firstly passed through pointwise convolution to change the channel dimension to 32. Subsequently, it undergoes depthwise convolution with 32 groups to extract features within each channel. Finally, pointwise convolution is applied again to fuse information between channels, resulting in an output image with 32 channels.

2.4.3. Vision Transformer Base/16

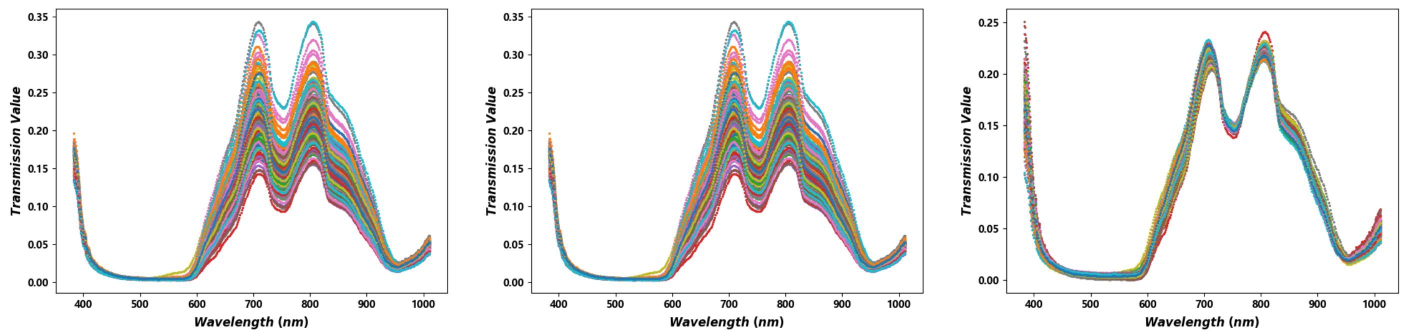
The backbone applied is Vision Transformer Base-16(ViT-16), which size of kernel stride is 16(Conv2d-16) in Patch Embedding Block. Generally, ViT-16 mainly consists of Patch Embedding, Position and Class Embedding, Encoder Layers, and MLP head. Firstly, the Patch Embedding layer divides the input image into patches of size 14*14 using Conv2d-16. The channel dimension is expanded from 32 to 768 through the convolution, resulting in an output image with dimensions of 768*14*14. Subsequently, with width and height dimensions flatten, it becomes 768*196(196 patches, each with a dimension of 768). Then a class token is added before the patches, and overall position embedding is added to incorporate positional information. The data is then fed into the Encoder Layers, consisting of 12 layers of Encoders, to extract relevant information. Finally, the class token for all patches is



(a) Spectrums on d 0, 2, 4, 6

(b) Spectrums on d 8, 10, 12, 14

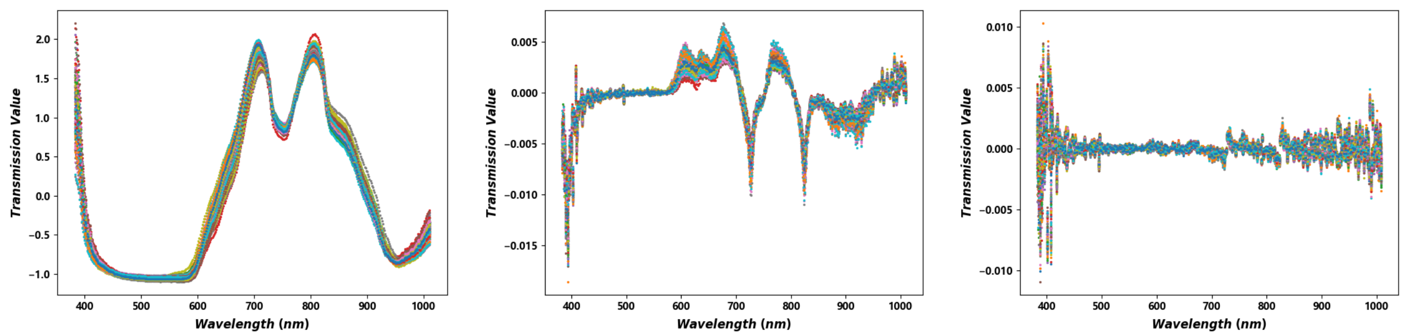
Figure 3: Average spectrum curve of male and female class. Channel image at each wavelength was represented as avergae value of all pixels.



(a) Original spectrums

(b) SG smoothed spectrums

(c) MSC smoothed spectrums



(d) SNV smoothed spectrums

(e) D1 smoothed spectrums

(f) D2 smoothed spectrums

Figure 4: The average spectrums of ROI smoothed by SG, MSC, SNV, D1, and D2 on d 10.

extracted, and after one Linear and Softmax in the MLP head, it is directly projected to the gender classes with probabilities, thus the in-ovo sexing by EggFormer is relized.

Due to the lack of inductive bias, a form of prior knowledge, in the transformer attention mechanism compared to CNN in ViT, the performance may be slightly inferior when the dataset is not large enough(Dosovitskiy et al., 2020). Therefore, we employed a transfer learning approach where the model weights, originally trained on the imageNet-21k dataset in the Encoder layers, were partially frozen. Throughout the model training process, several techniques were employed to achieve improved performance, including DropPath(Larsson et al., 2016), Cosine

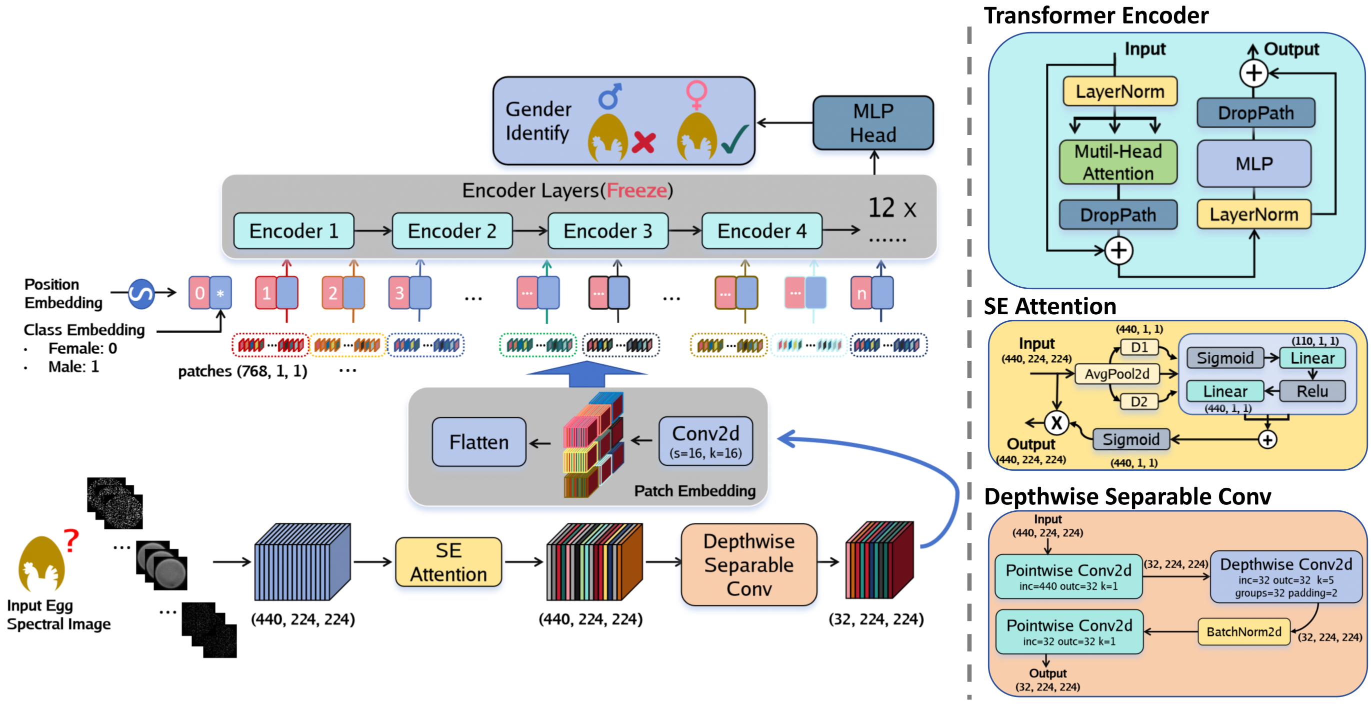


Figure 5: The structure of EggFormer(left) and the details in functional blocks(right).

Learning Rate Decay(He et al., 2019), and Kornia image augumentation(E. Riba and Bradski, 2018, 2020a,b).

2.5. Performance Evaluation Scores

To comprehensively evaluate the model built, indicators containing overall accuracy(OA), f1 score, Kappa coefficient were used. The calculating formulas were shown as follows. In the equations, TP(True Positive) is the number of samples that are correctly predicted to be positive cases, and TN(True Negative), FP(False Positive), FN(False Negative) can be referred by parity of reasoning. In models evaluation, OA means the proportion of samples that the model prediction results agree with the actual label, while AA stands for the average accuracy, which is the average proportion of correctly predicted cases for each class.

$$OA = \frac{TP + TN}{TP + FN + FP + TN} \tag{1}$$

$$AA = \frac{1}{2} * \frac{TP}{TP + FN} * \frac{TN}{TN + FP} \tag{2}$$

The Kappa coefficient takes into account the difference between the expected accuracy and the actual accuracy and is used to measure the consistency of the model classification, which is can be calculated by Equation(3-4).

$$Kappa = \frac{OA - PE}{1 - PE} \tag{3}$$

$$PE = \frac{(TN + FN) * (TN + FP) + (TP + FN) * (TP + FP)}{(TN + TP + FN + FP)^2} \tag{4}$$

The F1 score is the harmonic average of Precision and Recall, where can be calculated as Equation(5-7).

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

3. Results and Discussion

The proposed model was mostly implemented in Python, except for the CARS code, which is adopted from the original author and implemented in Matlab(Li et al., 2009). The work was implemented based on Pytorch and Scikit-learn open-source framework. The training and test platform hardware includes the Nvidia A800 GPU and AMD EPYC 7742 64-core processor, with 200G memory. The experiments were conducted using Leave-One-Out Cross-Validation (LOOCV), where the datasets were evenly divided into three parts. Two of these parts were allocated for training data(87 eggs), while the remaining one served as the test set(44 eggs), and the ratio of male and female in each part were kept as euqal as possible. Given the higher value of female chicks over male ones in practical poultry production, the model task was configured as binary classification, designating male eggs as positive samples and female eggs as negative ones. This approach aims to enhance the accuracy of male chick prediction, facilitating early screening. Subsequently, three models were obtained by 3 CVs, and the final scores were averaged as a comprehensive assessment of the method, and finnaly the optimal model was applied for interpretation and feature wavelength selection.

3.1. Significant Wavelengths Extracted by RF, SPA, PCA, and CARS

The average spectrum on even days from 0 to 14 were preprocessed by 5 smooth methods(including SG, MSC, SNV, D1, and D2) with RF. The result showed that all 5 methods achieved their maximum accuracy on day 10(Figure A.1). The accuracy of the original spectrums, as well as the D1 and D2, exceeds 0.9 on day 10, specifically reaching 0.939, 0.923, and 0.931, respectively. The ORI spectrum demonstrates superior performance in both accuracy, precision, f1, and Kappa scores.

Consequently, based on the best results obtained, we applied the original average spectrum on day 10 in the following significant wavelengths extract. The 4 dimensionality reduction methods (RF, PCA, SPA, CARS) all employed PLS-DA for classifying. As shown in Table 3, RF achieved the highest accuracy of 93.9%, while CARS performed the poorest with 80.8%. By interpretation of the 4 models, the top-10 most relevant wavelengths were selected and listed in Table 2. In terms of RF, it reduced the input 440 wavelengths to 219 (n_components=0.98), which accounts for 98% importance of all features. Among them, there were 9 wavelengths with contributions exceeding 2.5%, and contribution at 712.76nm was the highest with 4.42%. Specifically, the top-10 wavelengths were distributed from 698nm to 765nm, and the bands with contributions exceeding 0.05% were distributed in the range of 600-900nm, covering the visible and NIR regions(Figure A.2a). Besides, we also used the Shapley Additive exPlanations(SHAP) method(Shapley et al., 1953) to explain the best RF model in Figure A.2b. The each point represents one egg sample, and the higher of feature in Y-axis, more relevant is the wavelength. Specificlly, 712.76nm is the most relevant wavelength, and is positive related with male eggs when average value at this wavelength is high(red), and negative when value is low(blue). In general, the top-10 relevant wavelengths provided by SHAP fall within the range of 695-765nm.

Concerning PCA, the top-3 principal components (PCs) collectively accounted for over 99% of the wavelengths, with individual contributions of 95.57%, 3.04%, and 1.07%, respectively. By multiplying the values of the corresponding channels with the absolute values of the wavelength's eigenvalues, principal component maps for PC1, PC2, and PC3 were generated (Figure A.3). PC1 depicted the spectral image of the egg, and notably in Figure A.3a, an enlarged 5-pixel radius was applied during the ROI cropping to retain additional information, resulting in only a small ring of noise surrounding the egg. PC2 highlighted more pronounced noise, while the granular noise in PC3 was blended with the egg region, predominantly stemming from the camera scanning process. The Top-10 wavelengths for PC1 were distributed within the range of 702nm-714nm, with minimal differences in Eigenvectors.

The SPA algorithm identified 11 candidate wavelengths and finalized 10 as the selected wavelengths, while CARS featured 25 significant wavelengths. The wavelength ranges chosen by both 2 algorithms exhibited similarities, predominantly falling within 400-600nm, with a few extending to 800-1000nm. These wavelengths were marked as scattered points on the average spectrum in Figure A.4a and A.4c. Additionally, the changing log of scores during the model calculation iterations were recorded separately in Figure A.4b and A.4d.

3.2. Performance of Models

3.2.1. Significant wavelengths performance

Due to the limited number of the training dataset, the Kornia library(Riba et al., 2020) was employed for data augmentation during model training. This included RandomHorizontalFlip, RandomRotation(±180°), and RandomErasing(scale:0.02-0.05, ratio:0.3-3.3), all random probabilities were set to 0.8. The experimental results are presented in Table3, and to validate the effectiveness of reduction algorithms, channel images of full wavelengths and recoupled with selected wavelengths were chosen as inputs. When employing the ViT-Base model, the overall accuracy using all bands was the lowest with 0.924, indicating shortcomings in the ViT model when handling of high-dimensional data. Despite SPA retaining fewer bands(10) compared to CARS(25) after bands selection, its accuracy was worse, suggesting potential loss of essential information during dimensionality reduction. Notably, CARS exhibited advantages in both feature wavelength selection and accuracy, achieving the highest accuracy of 0.939. Besides, PCA-PC1 selected 267 wavelengths, accounting for 98% of the variance in PC1, with comparable performance to RF. These observations were consistent when utilizing models of ViT-Img21K and EggFormer.

3.2.2. Comparison of models

In order to evaluate the performance of models, 3 models were compared: the ViT-Base model, ViT-Base-Img21K with pre-trained weights, and proposed EggFormer. Considering the differences in models, SGD and AdamW optimizer were both used to get the best results, with the batch size of 64, epochs of 120, leraning rate of 5e-4, and the finnal results were shown in Table3. Initially, without pre-trained weights, the accuracy of ViT only matched with RF, but significantly improved when pre-trained weights were applied, reaching an accuracy of 0.939 with full-band input. However, this came

Table 2: The top-10 significant wavelengths by RF, PCA, SPA, and CARS.

Rank	**RF**		**PCA-PC1**		**SPA**		**CARS**	
	nm	**Contribution(%)**	**nm**	**Eigenvectors**	**nm**	**Relevance**	**nm**	**Coefficients**
1	712.76	4.42	706.94	0.095	810.73	0.36	402.67	1.88
2	714.22	4.10	708.39	0.095	515.08	0.29	413.41	1.71
3	765.33	3.96	705.48	0.095	522.09	0.21	448.68	1.38
4	722.97	3.68	709.85	0.095	406.69	0.12	481.69	1.35
5	763.87	3.22	704.03	0.094	432.34	0.10	522.09	1.33
6	720.05	2.76	711.31	0.094	571.51	0.07	439.13	1.28
7	717.13	2.66	702.57	0.094	408.03	0.07	424.21	1.28
8	702.57	2.52	712.76	0.094	994.74	0.06	429.62	1.26
9	718.59	2.51	701.12	0.094	495.56	0.02	408.03	1.25
10	698.21	2.37	714.22	0.093	936.61	0.01	390.65	1.22

Table 3: The scores of models. The number in brackets are number of significant wavelengths extracted by the corresponding method.

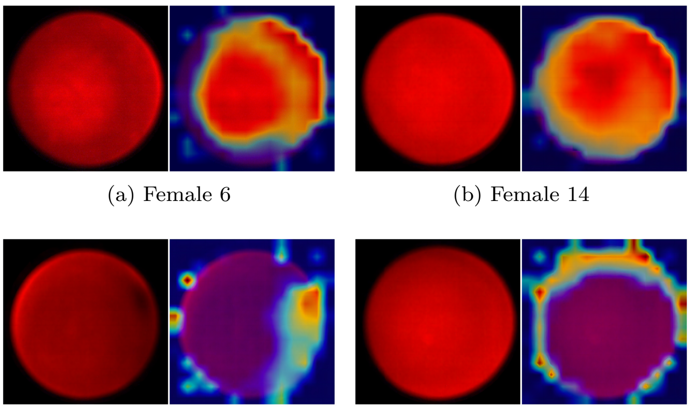
Model	**Input bands**	**Accuracy**	**Precision**	**Recall**	**F1 score**	**Kappa**	**Train Params**	**Total Params**
RF	All(440)	0.939	0.931	0.957	0.943	0.877	\	\
PCA-PLSDA	All(440)	0.901	0.870	0.957	0.911	0.799	\	\
SPA-PLSDA	All(440)	0.839	0.795	0.943	0.863	0.673	\	\
CARS-PLSDA	All(440)	0.808	0.771	0.914	0.837	0.609	\	\
ViT-Base	All(440)	0.924	0.943	0.913	0.926	0.847	171.57M	171.57M
	RF(219)	0.931	0.943	0.928	0.935	0.862	128.12M	128.12M
	PCA-PC1(267)	0.931	0.947	0.928	0.934	0.862	137.55M	137.55M
	SPA(10)	0.924	**0.958**	0.899	0.925	0.044	87.02M	**87.02M**
	CARS(25)	0.939	0.908	**0.986**	0.945	0.876	89.97M	89.97M
ViT-Base-Img21K	All(440)	0.939	0.948	0.942	0.941	0.877	86.51M	171.57M
	RF(219)	0.939	0.932	0.957	0.942	0.877	43.06M	128.12M
	PCA-PC1(267)	0.939	0.932	0.957	0.942	0.877	128.12M	137.55M
	SPA(10)	0.931	0.931	0.928	0.934	0.862	**1.97M**	**87.02M**
	CARS(25)	0.946	0.933	0.971	0.951	0.892	4.92M	89.97M
EggFormer	All(440)	**0.954**	0.933	**0.986**	**0.958**	**0.908**	6.60M	91.66M
	RF(219)	0.946	0.933	0.971	0.951	0.892	6.37M	91.43M
	PCA-PC1(267)	0.946	0.933	0.971	0.951	0.892	6.41M	91.47M
	SPA(10)	0.916	0.957	0.885	0.918	0.832	6.30M	91.35M
	CARS(25)	**0.954**	0.933	**0.986**	**0.958**	**0.908**	6.30M	91.35M

with a notable increase in number of parameters. Featuring SE channel attention and DWConv, EggFormer achieved the best accuracy of 0.954 with full-band input, surpassing all models in both f1 and Kappa scores, which representing comprehensive evaluation results. Furthermore, EggFormer demonstrated a parameter advantage over ViT models, with approximately 6.5M training parameters and 91.5M total parameters. Interestingly, after spectral dimensions reduced by RF, PCA, and CARS methods, EggFormer also outperformed other models, indicating its ability to extract latent information from spectral images. However, the performance of EggFormer after SPA was lower than ViT-Base-21K, possibly due to the initial reduction in bands by 3/4 with SE channel attention, which limits accuracy.

3.2.3. Interpretation of EggFormer

The training and validation log curves of 3 cross validations were shown in FigureA.5. With the increasing number of epochs, the training and testing loss decreased to approximately 0.25, with a high confidence probability. To further clear the internal working mechanisms of EggFormer, we conducted visualizations by Grad-CAM(Selvaraju et al., 2017), based on gradients. As depicted, heatmap visualization allows us to comprehend the focused areas by model, with increasing red color intensity indicating higher attention levels. When using the last layer's norm1 block of the Encoder Layers as anaylsis target, as exemplified by the both 2 female and male eggs, EggFormer's attention is primarily directed towards the interior or edges of the eggs. Notably, when the egg is classified as female, EggFormer tends to emphasize brighter regions of the egg. Conversely, for male classification, attention shifts towards the edges or darker regions within the egg. This observation is also supported by the average spectrums of the eggs in Figure3, wherein the peak value of male eggs spectrums are indeed lower than those of female ones.

To dissect the SE channel mechanism of EggFormer, the output layer of SE is extracted and normalized, as illustrated in the Figure7. The wavelength contributions shown in the figure are distributed across the entire spectrum. To reduce the number of effective wavelengths, a sampling interval is set, with unique values taken between intervals based on decreasing contribution values. Experimental results were depicted in the figure when the learning rate is set to 6e-4 and the number of channels is less than or equal to 10, with a subsampling ratio of 1 for SE channels. When intervals are set as factors of 440(like 1, 2, 4, 5, 10, 20, 22), a 4-wavelengths input comprising 945.354nm, 762.407nm, 409.377nm, and 644.587nm maintains an accuracy of 0.946. Simultaneously, even with a 22-wavelengths input(less than CARS), EggFormer sustains an accuracy of 0.954, with corresponding wavelengths indicated in the figure caption and marked on the figure. The reduction of bands contributes to enhancing recognition efficiency, lowering costs, and consequently increasing economic returns in industrial applications.



(a) Female 6 (b) Female 14

(c) Male 72 (d) Male 174

Figure 6: RGB and Heapmap by Grad-CAM of eggs on d 10.

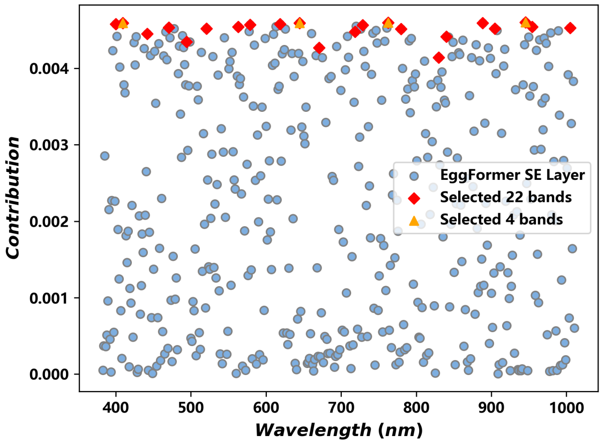


Figure 7: Contributions of wavelengths by EggFormer. The selected 4 bands are 945.354, 762.407, 409.377, and 644.587. The slected 22 bands are as follows: 945.354, 762.407, 409.377, 644.587, 888.39, 399.991, 618.669, 578.621, 728.801, 954.091, 562.989, 470.638, 1004.87, 904.482, 520.684, 779.97, 718.592, 441.857, 840.049, 494.166, 670.619, and 829.789.

3.2.4. EggFormer on d 0-14

The results of EggFormer using full-bands input on even-number days were shown in Figure8. The accuracy peaks on the 10th day, with a trend similar to the RF prediction results shown in the previous figure: accuracy gradually increases from day 0 to 10 before declining. This phenomenon occurs because during the early stages of egg development, when embryos are just beginning to form, the differences between male and female embryos are primarily at the genetic level. Organs such as the gonads, which are associated with gender, have not yet developed. Around d 6.5, gonadal differentiation begins(Hirst et al., 2018). Subsequently, as the embryo develops feathers and organs, interference signals increase, reducing egg transparency. The hormones of both sexes can be measured in the allantoic fluid on day 7 to 14(Kaleta and Redmann, 2008, Gill et al., 1983, Phelps et al., 2003). During the development process, the content of hemoglobin in the blood varies according to gender(Galli et al., 2016), with band around 785nm, consistent with 762.407nm and 779.97 nm in the selected bands.

4. Conclusions

In our study, EggFormer was developed and realized for in-ovo sexing of Hy-Line Sonia eggs with hyperspectral imaging, which utilizes ViT-Base as backbone. Hyperspectral images were collected on even days form 0 to 14, and compared with conventional methods for selecting significant wavelengths. EggFormer possessed the capacity of handling the full-bands images input, and chieved the highest accuracy of 0.954, an f1 score of 0.958, and a Kappa of 0.908 on day 10. Besides, the working mechanism of EggFormer was interpreted and subsequently reduced the full set of 440 bands to 22 bands while maintaining the same accuracy of 0.954, and to 4 bands with an accuracy of 0.946. EggFormer demonstrates the capability to extract latent information from spectral images, offering promise for early-stage embryo sex identification using fewer wavelengths, with potential applications in hatch industry. However, the limited size of our dataset constrains the maximum recognition accuracy of deep learning models. We anticipate that gathering more spectral images in future applications will enable more precise identification.

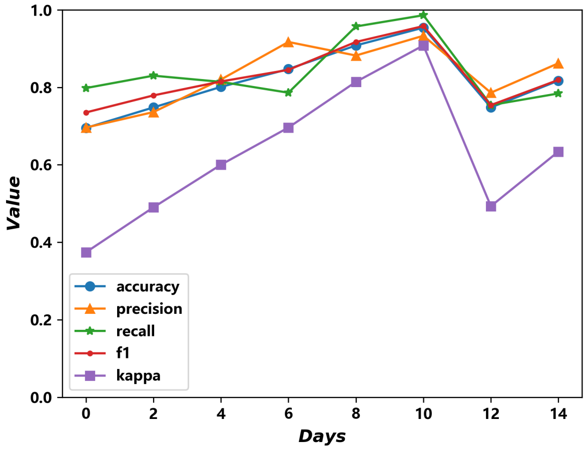


Figure 8: Figure results of EggFormer on even days 0-14

5. Acknowledgement

We gratefully acknowledge the College of Food Science and Technology of Naingjing Agricultural University for providing the full hyperspectral image datasets of Hy-Line Sonia. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

6. Data availability

The data has to be used further and cannot be shared for the time being.

7. Authors' contributions

JCM and HJX: Conceptualization, Formal analysis, Investigation, Methodology, Data curation, Model building, Visualization, Writing-original draft preparation, Writing-reviewing and Editing. **SK:** Data curation, Conceptualization. **CZX:** Investigation, Methodology, Model building. **WXY:** Investigation, Methodology.

XHL, HJX, TK, PLQ: Supervision, Project administration, Funding acquisition. All authors read and approved the final manuscript.

References

M.-E. Krautwald-Junghanns, K. Cramer, B. Fischer, A. Förster, R. Galli, F. Kremer, E. Mapesa, S. Meissner, R. Preisinger, G. Preusse, et al., Current approaches to avoid the culling of day-old male chicks in the layer industry, with special reference to spectroscopic methods, Poultry science 97 (2018) 749–757.

N. Jia, B. Li, J. Zhu, H. Wang, Y. Zhao, W. Zhao, A review of key techniques for in ovo sexing of chicken eggs, Agriculture 13 (2023) 677.

L. He, P. Martins, J. Huguenin, T.-N.-N. Van, T. Manso, T. Galindo, F. Gregoire, L. Catherinot, F. Molina, J. Espeut, Simple, sensitive and robust chicken specific sexing assays, compliant with large scale analysis, PloS one 14 (2019) e0213033.

A. Di Concetto, O. Morice, M. Corion, S. Monteiro Belo dos Santos, Chick and duckling killing: Achieving an eu-wide prohibition, Chick and Duckling Killing: Achieving an EU-Wide Prohibition (2023).

M. Rosenbruch, [the sensitivity of chicken embryos in incubated eggs][article in german], ALTEX-Alternatives to animal experimentation 14 (1997) 111–113.

M. Corion, S. Santos, B. D. Ketelaere, D. Spasic, M. Hertog, J. Lammertyn, Trends in in ovo sexing technologies: insights and interpretation from papers and patents, 2023.

D. J. Mellor, T. J. Diesch, Birth and hatching: key events in the onset of awareness in the lamb and chick, New Zealand veterinary journal 55 (2007) 51–60.

M. Corion, J. Keresztes, B. D. Ketelaere, W. Saeys, In ovo sexing of eggs from brown breeds with a gender-specific color using visible-near-infrared spectroscopy: effect of incubation day and measurement configuration, Poultry Science 101 (2022).

R. Galli, G. Preusse, O. Uckermann, T. Bartels, M. E. Krautwald-Junghanns, E. Koch, G. Steiner, In ovo sexing of domestic chicken eggs by raman spectroscopy, Analytical Chemistry 88 (2016) 8657–8663.

N. Jia, B. Li, J. Zhu, H. Wang, Y. Zhao, W. Zhao, A review of key techniques for in ovo sexing of chicken eggs, 2023.

M. Corion, S. Santos, B. S. Ketelaere, D. Spasic, M. Hertog, J. Lammertyn, Insights and interpretation of the trends for in ovo sexing technologies in papers and patents (2023).

G. Hu, Y.-g. Jin, G.-f. Jin, M.-h. Ma, et al., Nondestructive characterization gender of chicken eggs by odor using spme/gc-ms coupled with chemometrics, Poultry Science 101 (2022) 101619.

D. Göhler, B. Fischer, S. Meissner, In-ovo sexing of 14-day-old chicken embryos by pattern analysis in hyperspectral images (vis/nir spectra): A non-destructive method for layer lines with gender-specific down feather color, Poultry science 96 (2017) 1–4.

L. Pan, W. Zhang, Y. Sun, X. Gu, L. Ma, Z. Li, P. Hu, K. Tu, et al., Gender determination of early chicken hatching eggs embryos by hyperspectral imaging, Transactions of the Chinese Society of Agricultural Engineering 32 (2016) 181–186.

A. Rahman, A. Khaliduzzaman, T. Suzuki, N. Kondo, Non-destructive technologies for embryo gender prediction, in: Informatics in Poultry Production: A Technical Guidebook for Egg and Poultry Education, Research and Industry, Springer, 2022, pp. 77–95.

J. Dong, X. Dong, Y. Li, Y. Peng, K. Chao, C. Gao, X. Tang, Identification of unfertilized duck eggs before hatching using visible/near infrared transmittance spectroscopy, Computers and Electronics in Agriculture 157 (2019) 471–478.

M. Ghaderi, S. A. Mireei, A. Masoumi, M. Sedghi, M. Nazeri, Fertility detection of unincubated chicken eggs by hyperspectral transmission imaging in the vis-swnir region, Scientific Reports 14 (2024) 1289.

C. Toksoz, M. Albayrak, H. Yasar, Chicken egg sexing by using data mining process, Fresenius Environ. Bull 30 (2021) 1373–1381.

R. Galli, G. Preusse, O. Uckermann, T. Bartels, M. E. Krautwald-Junghanns, E. Koch, G. Steiner, In ovo sexing of chicken eggs by fluorescence spectroscopy, Analytical and Bioanalytical Chemistry 409 (2017) 1185–1194.

P. Horkaew, S. Kupittayanant, P. Kupittayanant, Non-invasive in ovo sexing in korat chicken by pattern recognition of its embryologic vasculature, Journal of Applied Poultry Research (2024) 100424.

N. Jia, B. Li, Y. Zhao, S. Fan, J. Zhu, H. Wang, W. Zhao, Exploratory study of sex identification for chicken embryos based on blood vessel images and deep learning, Agriculture 13 (2023) 1480.

A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An image is worth 16x16 words: Transformers for image recognition at scale (2020).

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Advances in neural information processing systems 30 (2017).

Y. Cao, P. Yuan, H. Xu, J. F. Martínez-Ortega, J. Feng, Z. Zhai, Detecting asymptomatic infections of rice bacterial leaf blight using hyperspectral imaging and 3-dimensional convolutional neural network with spectral dilated convolution, Frontiers in Plant Science 13 (2022) 963170.

M. Belgiu, L. Drăguţ, Random forest in remote sensing: A review of applications and future directions, ISPRS journal of photogrammetry and remote sensing 114 (2016) 24–31.

M. A. Shahin, S. J. Symons, Detection of fusarium damaged kernels in canada western red spring wheat using visible/near-infrared hyperspectral imaging and principal component analysis, Computers and electronics in agriculture 75 (2011) 107–112.

C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, A. Zeileis, Conditional variable importance for random forests, BMC bioinformatics 9 (2008) 1–11.

L. Cao, K. S. Chua, W. Chong, H. Lee, Q. Gu, A comparison of pca, kpca and ica for dimensionality reduction in support vector machine, Neurocomputing 55 (2003) 321–336.

B. M. S. Hasan, A. M. Abdulazeez, A review of principal component analysis algorithm for dimensionality reduction, Journal of Soft Computing and Data Mining 2 (2021) 20–30.

S. F. C. Soares, A. A. Gomes, M. C. U. Araujo, A. R. Galvão Filho, R. K. H. Galvão, The successive projections algorithm, TrAC Trends in Analytical Chemistry 42 (2013) 84–98.

J. Sun, X. Zhou, Y. Hu, X. Wu, X. Zhang, P. Wang, Visualizing distribution of moisture content in tea leaves using optimization algorithms and nir hyperspectral imaging, Computers and Electronics in Agriculture 160 (2019) 153–159.

H. Li, Y. Liang, Q. Xu, D. Cao, Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration, Analytica chimica acta 648 (2009) 77–84.

J. Liu, S. Liu, S. Shin, F. Liu, T. Shi, C. Lv, Q. Qiao, H. Fang, W. Jiang, H. Men, Detection of apple taste information using model based on hyperspectral imaging and electronic tongue data., Sensors & Materials 32 (2020).

W. Ai, S. Liu, H. Liao, J. Du, Y. Cai, C. Liao, H. Shi, Y. Lin, M. Junaid, X. Yue, et al., Application of hyperspectral imaging technology in the rapid identification of microplastics in farmland soil, Science of The Total Environment 807 (2022) 151030.

J. Ma, D.-W. Sun, H. Pu, Spectral absorption index in hyperspectral image analysis for predicting moisture contents in pork longissimus dorsi muscles, Food chemistry 197 (2016) 848–854.

Q. Guo, W. Wu, D. Massart, The robust normal variate transform for pattern recognition with near-infrared data, Analytica chimica acta 382 (1999) 87–103.

A. Femenias, M. B. Bainotti, F. Gatius, A. J. Ramos, S. Marín, Standardization of near infrared hyperspectral imaging for wheat single kernel sorting according to deoxynivalenol level, Food Research International 139 (2021) 109925.

J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7132–7141.

C. Fran, et al., Deep learning with depth wise separable convolutions, in: IEEE conference on computer vision and pattern recognition (CVPR), 2017.

G. Larsson, M. Maire, G. Shakhnarovich, Fractalnet: Ultra-deep neural networks without residuals, arXiv preprint arXiv:1605.07648 (2016).

T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, M. Li, Bag of tricks for image classification with convolutional neural networks, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 558–567.

W. C. E. R. E. Riba, M. Fathollahi, G. Bradski, torch-geometry: when pytorch meets geometry, 2018. URL: https://drive.google.com/file/d/1xiao1Xj9WzjJ08YY_nYwsthE-wxfyfhG/view?usp=sharing.

J. S. D. P. F. M.-N. E. Riba, D. Mishkin, G. Bradski, A survey on kornia: an open source differentiable computer vision library for pytorch, 2020a.

D. P. E. R. E. Riba, D. Mishkin, G. Bradski, Kornia: an open source differentiable computer vision library for pytorch, in: Winter Conference on Applications of Computer Vision, 2020b. URL: https://arxiv.org/pdf/1910.02190.pdf.

L. S. Shapley, et al., A value for n-person games (1953).

E. Riba, D. Mishkin, D. Ponsa, E. Rublee, G. Bradski, Kornia: an open source differentiable computer vision library for pytorch, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020, pp. 3674–3683.

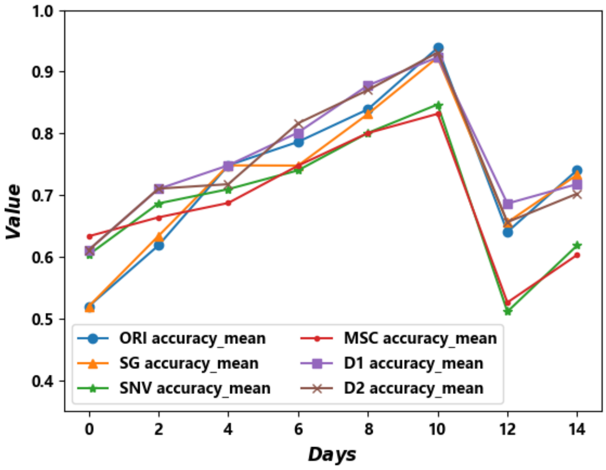
R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: Visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 618–626.

C. E. Hirst, A. T. Major, C. A. Smith, Sex determination and gonadal sex differentiation in the chicken model, International Journal of Developmental Biology 62 (2018) 153–166.

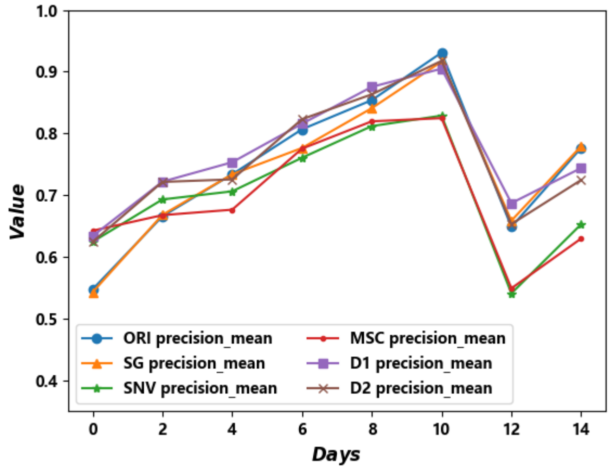
E. Kaleta, T. Redmann, Approaches to determine the sex prior to and after incubation of chicken eggs and of day-old chicks, World's Poultry Science Journal 64 (2008) 391–399.

D. V. Gill, H. A. Robertson, T. Betz, In vivo estrogen synthesis by the developing chicken (gallus gallus) embryo, General and Comparative Endocrinology 49 (1983) 176–186.

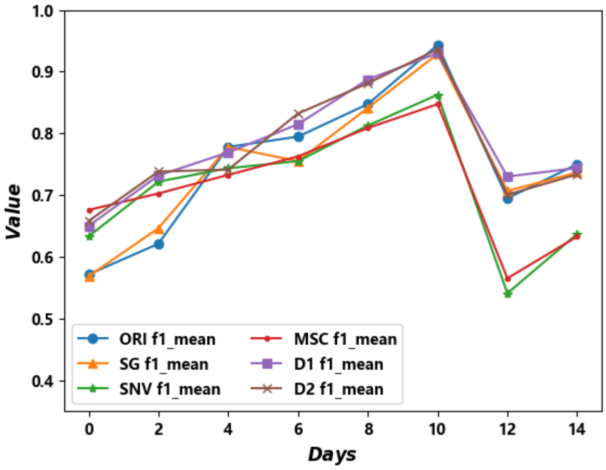
P. Phelps, A. Bhutada, S. Bryan, A. Chalker, B. Ferrell, S. Neuman, C. Ricks, H. Tran, T. Butt, Automated identification of male layer chicks prior to hatch, Worlds Poult Sci J 59 (2003) 33–8.



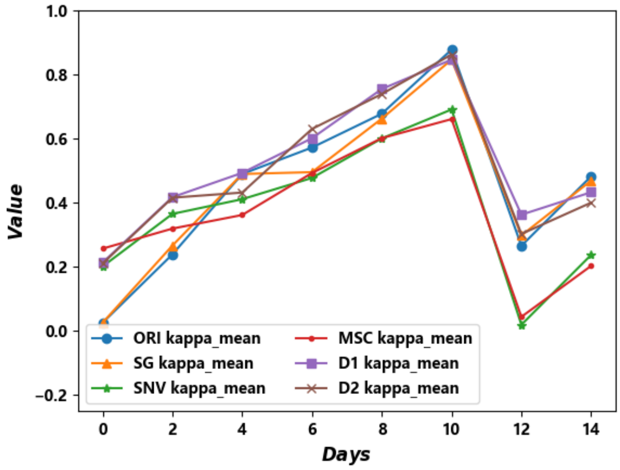
(a) Accuracy score



(b) Precision score

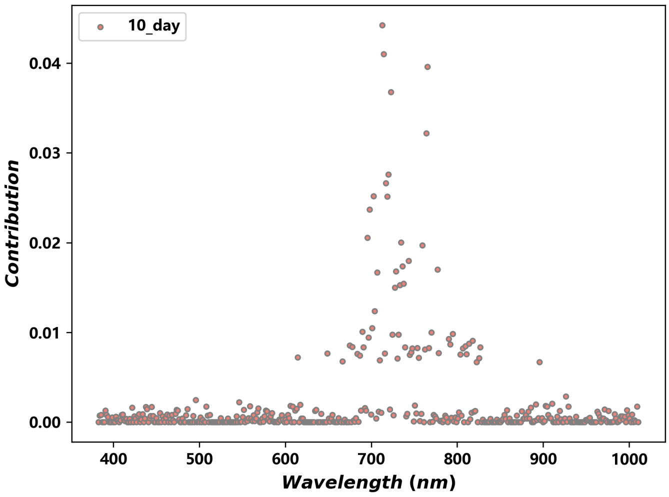


(c) F1 score

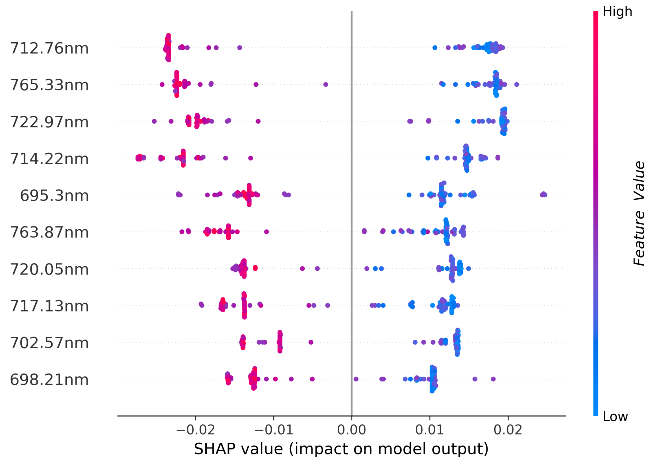


(d) Kappa score

Figure A.1: Scores comparing with 6 prepossing methods of average spectrum on d 0-14.

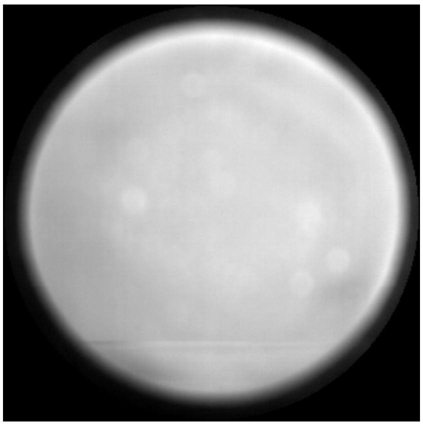


(a) Feature importances by best RF model

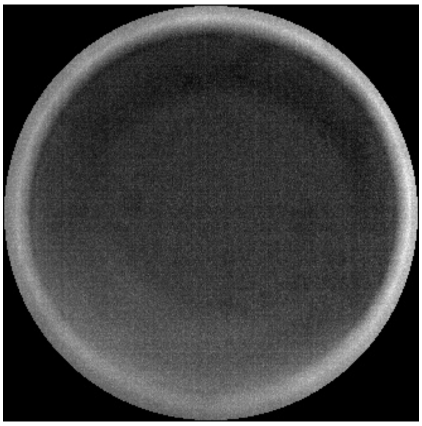


(b) Feature importances by SHAP on best RF model

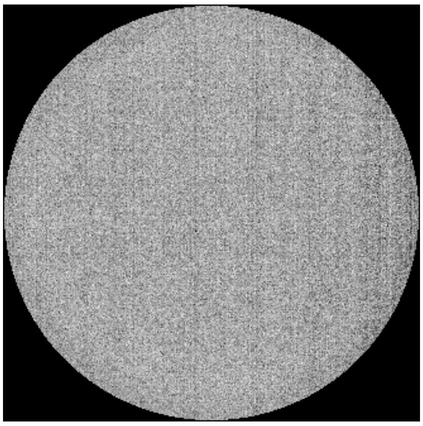
Figure A.2: Contributions of all wavelengths by RF on d 10



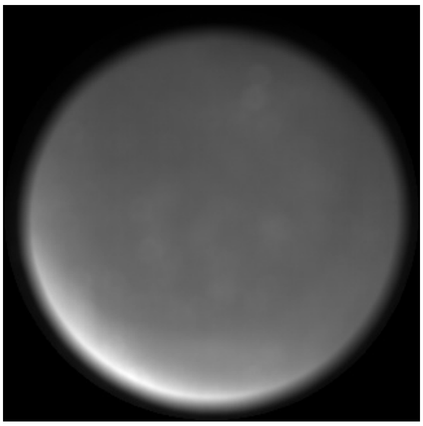
(a) PC1 of Male-1



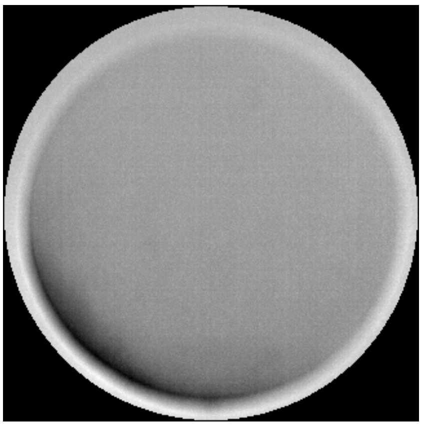
(b) PC2 of Male-1



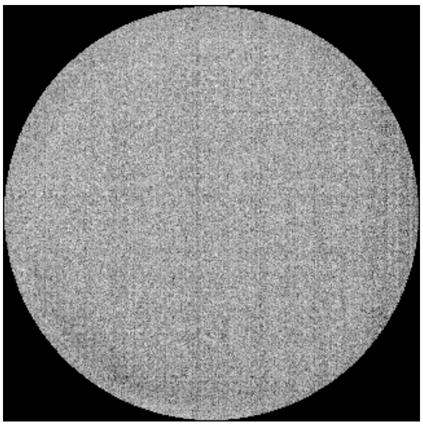
(c) PC3 of Male-1



(d) PC1 of Female-3

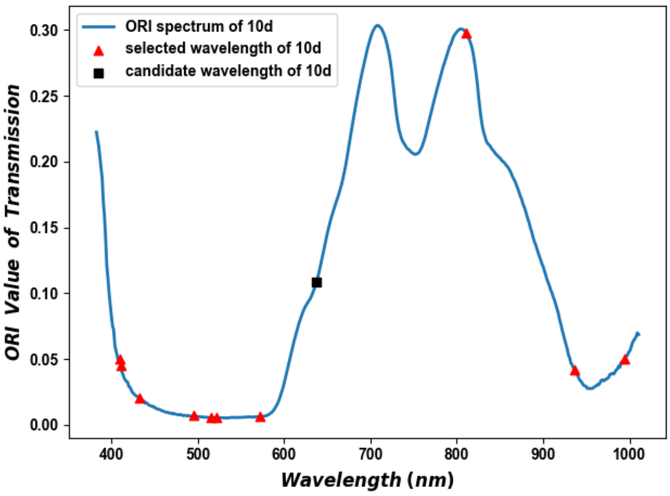


(e) PC2 of Female-3

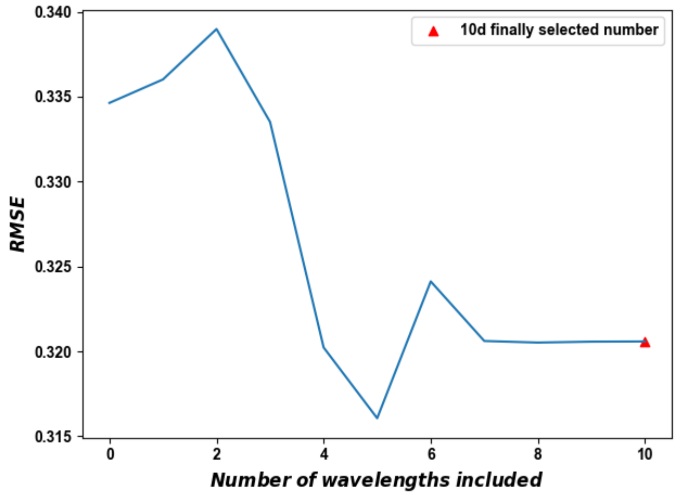


(f) PC3 of Female-3

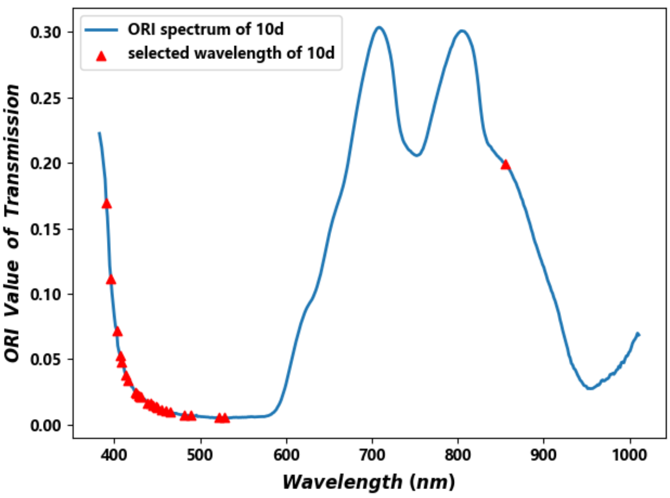
Figure A.3: Figure of PCs of male and female egg by PCA on d 10



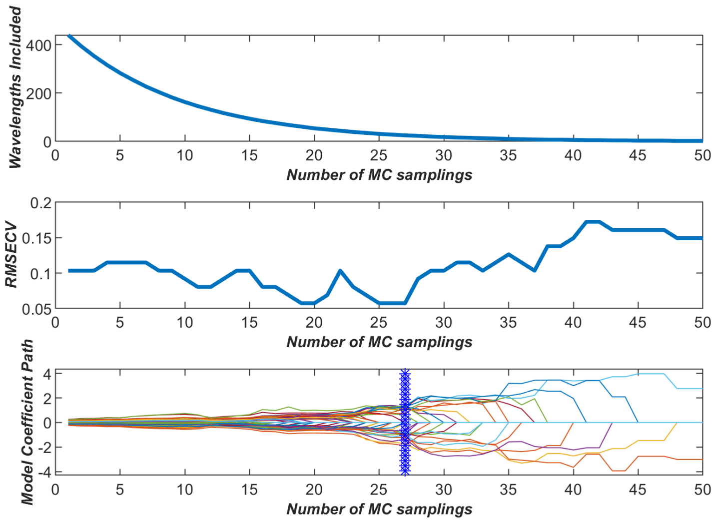
(a) Wavelengths selected by SPA



(b) RMSE of SPA

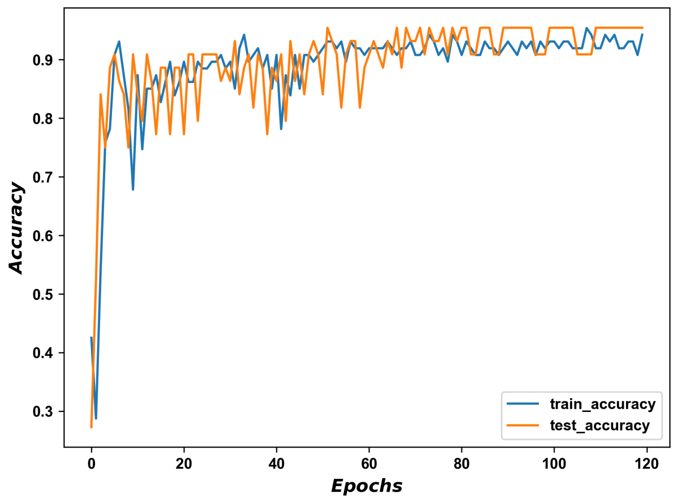


(c) Wavelengths selected by CARS

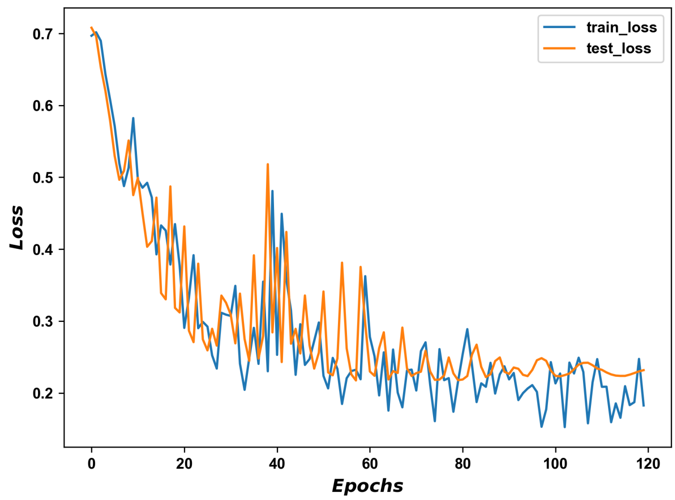


(d) Scores curve during CARS. Plots top and center show the changing of the number of sampled wavelengths and 5-fold RMSECV values. Plot bottom records the regression coefficient path of each wavelength. The vertical asterisk line denotes the optimal point where 5-fold CV values achieve the lowest.

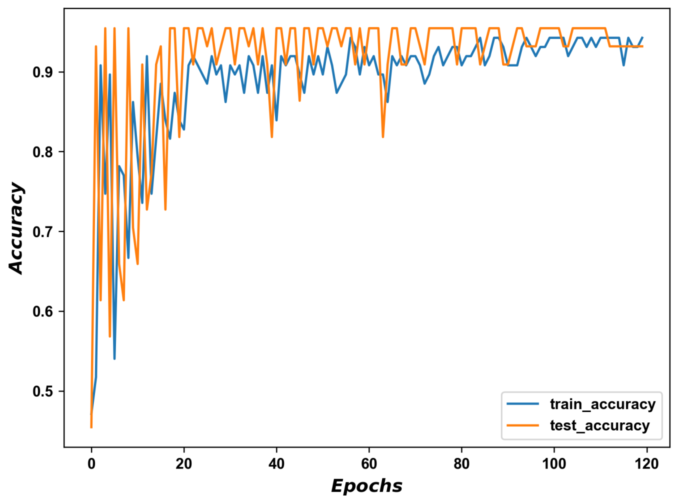
Figure A.4: Figure of slection of significant wavlengths by SPA and CARS on d 10.



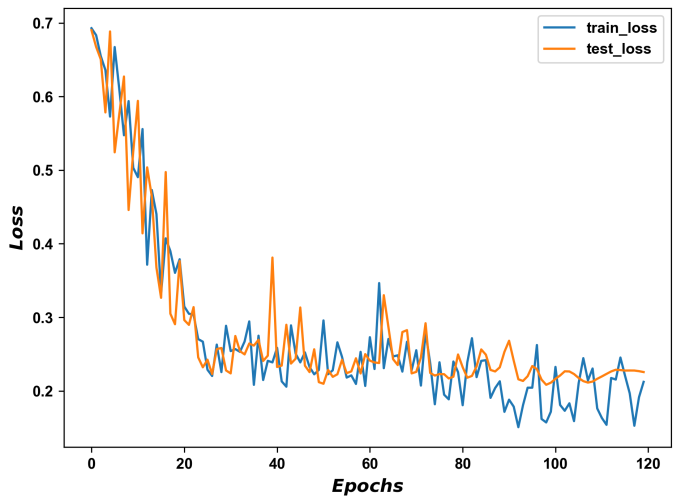
(a) Accuracy curve of cross validation 0



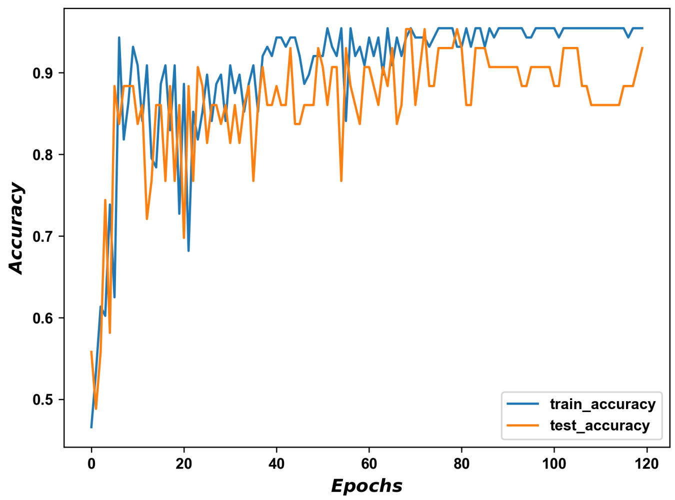
(b) Loss curve of cross validation 0



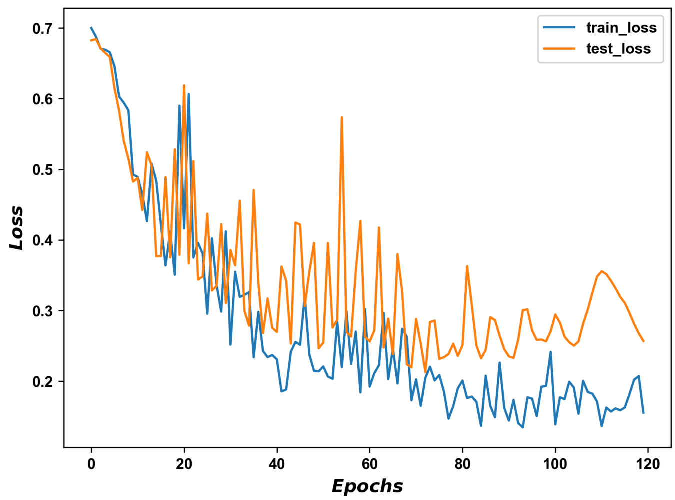
(c) Accuracy curve of cross validation 1



(d) Loss curve of cross validation 1



(e) Accuracy curve of cross validation 2



(f) Loss curve of cross validation 2

Figure A.5: Figure of accuracy and loss curve in 3 cross validations on d 10.