Technical University of Lodz
Institute of Electronics

University of Warmia and Mazury in Olsztyn

Computer vision quality assessment of barley kernels

Izabella Korczyńska[1], Aleksandra Jarosik[1], Piotr M. Szczypiński[1], Piotr Zapotoczny[2]

[1] Institute of Electronics, Lodz University of Technology
[2] Department of Agri-Food Process Engineering, University of Warmia and Mazury in Olsztyn

EFITA/WCCA/CIGR Conference, Poznań 2015

Supported by the National Centre for Research and Development in Poland,
grant no. PBS3/A8/38/2015

Visual Inspection and Computer Vision

1) Visual inspection is one of the oldest and reliable food quality assessment methods. However, in industrial environment it is labor-expensive.
2) With the development of computer vision algorithms the human-expert can be assisted or replaced by an automatic expert systems.
3) Coherent and optimal methodology for developing such systems is not yet established. Nevertheless each case can be solved individually using unique, tailored algorithms.

Quality expectations

In the brewing industry, crucial
for the production of high quality malt is:

- kernel uniformity,
- existence of germ in every kernel,
- lack of mechanical defects and
- lack of fungal infections

Scope of the Presentation

Goal and motivation

Methods

- Material
- Image segmentation
- Feature extraction
- Machine learning
- Data classification
- Cross-validation

Results

- Classification performance

Conclusions

Goal and motivation

Photo: Dan DeLong/Seattle Post-Intelligencer

Economic

- Inspect more
- Make inspection quick
- Simplify training

Technical

- Find reliable method
- Improve reproducibility

Material

Images acquired with flatbed scanner.

Samples of labeled kernels from malt ho-
use – *Soufflet Agro Polska.*

Classes:

- Good for malting
- Undeveloped
- Cracked
- Sprouted
- Swollen
- Missing germ
- Fusarium infected
- Other fungal infections

Sprouted

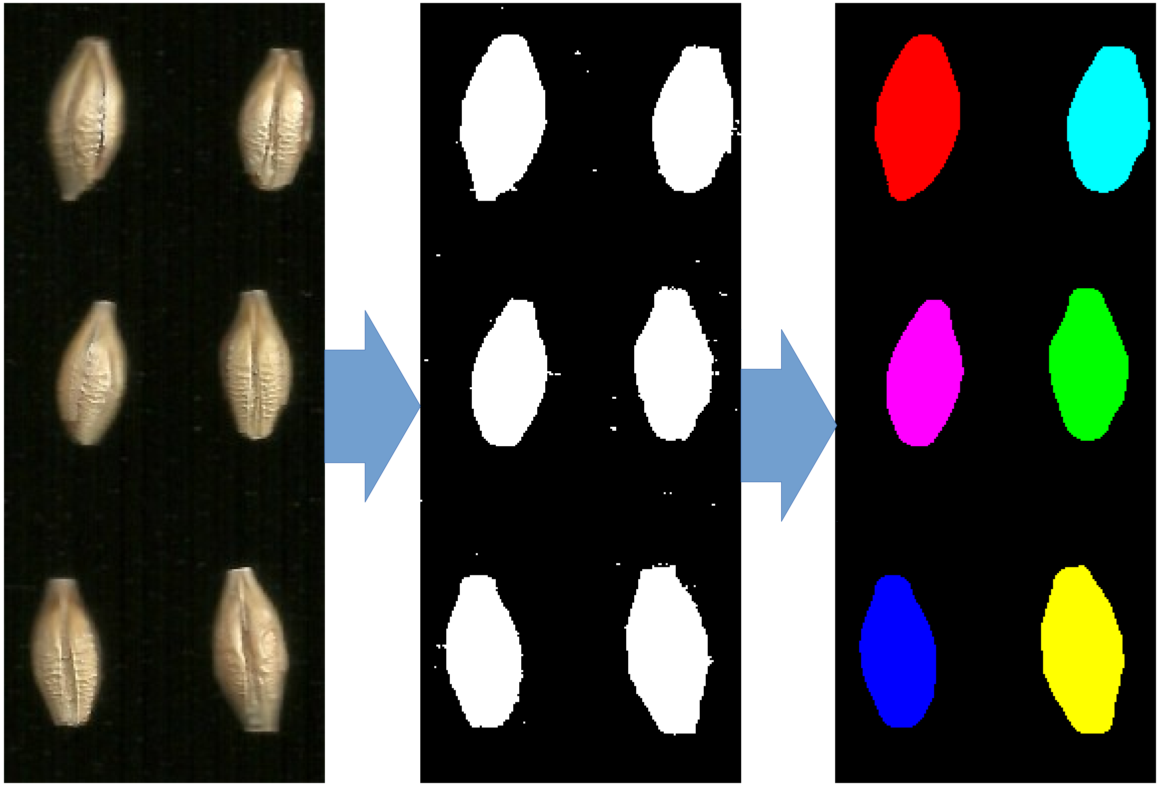
Fusarium infected

Image segmentation

Image segmentation is to partition an image into regions of interest suitable for further analysis. Here, the regions may represent individual kernels.

Our approach:

- Grey-level thresholding
- Median filtering
- Identifying disjoint areas

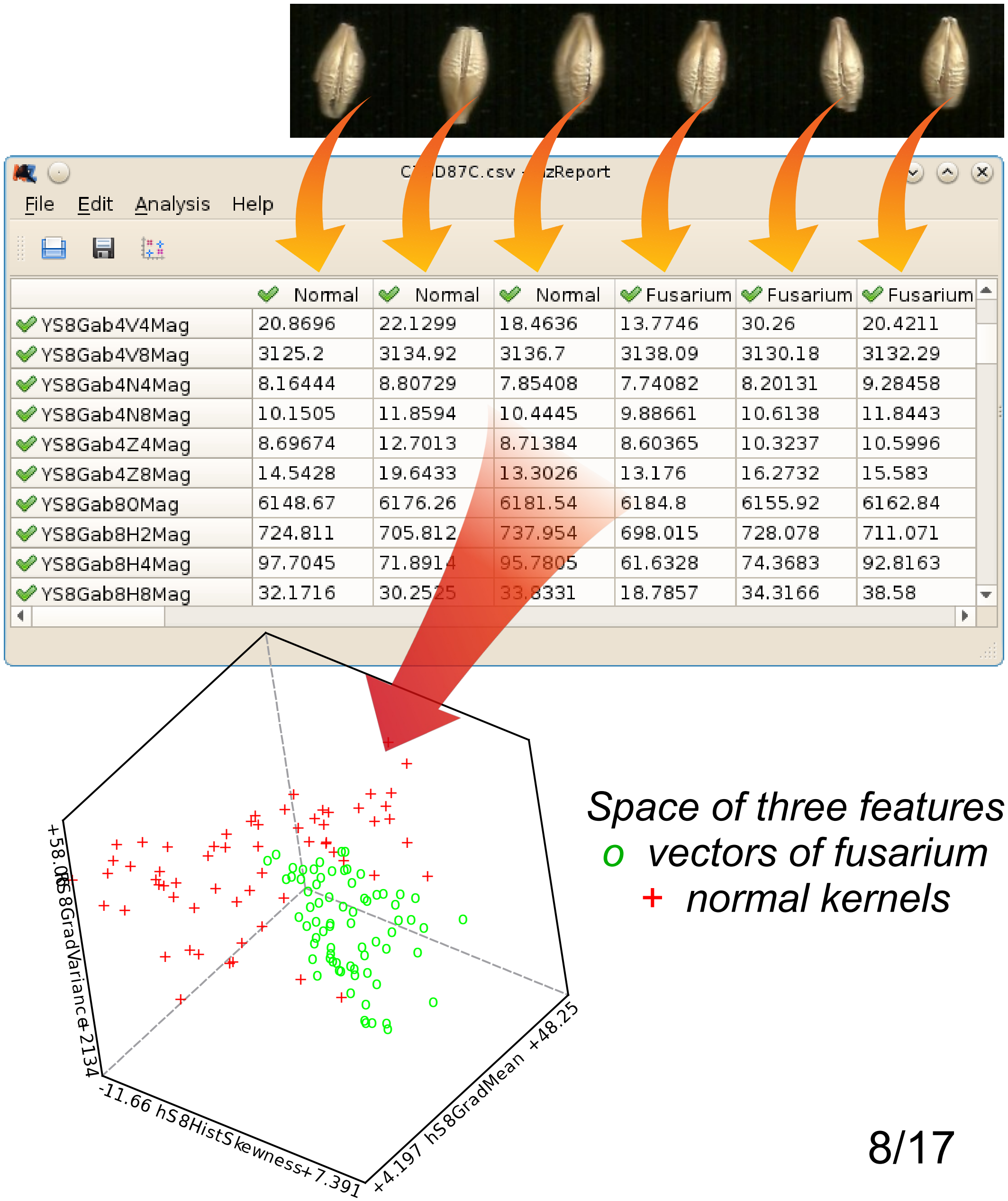


Feature extraction

Every kernel is characterized by numerical variables – features.

Such features quantitatively describe shape, texture and color.

Every kernel described by a vector of several-hundred features can be viewed as a coordinate in a several-hundred-dimensional space.



	Normal	Normal	Normal	Fusarium	Fusarium	Fusarium
YS8Gab4V4Mag	20.8696	22.1299	18.4636	13.7746	30.26	20.4211
YS8Gab4V8Mag	3125.2	3134.92	3136.7	3138.09	3130.18	3132.29
YS8Gab4N4Mag	8.16444	8.80729	7.85408	7.74082	8.20131	9.28458
YS8Gab4N8Mag	10.1505	11.8594	10.4445	9.88661	10.6138	11.8443
YS8Gab4Z4Mag	8.69674	12.7013	8.71384	8.60365	10.3237	10.5996
YS8Gab4Z8Mag	14.5428	19.6433	13.3026	13.176	16.2732	15.583
YS8Gab8OMag	6148.67	6176.26	6181.54	6184.8	6155.92	6162.84
YS8Gab8H2Mag	724.811	705.812	737.954	698.015	728.078	711.071
YS8Gab8H4Mag	97.7045	71.8914	95.7805	61.6328	74.3683	92.8163
YS8Gab8H8Mag	32.1716	30.2525	33.8331	18.7857	34.3166	38.58

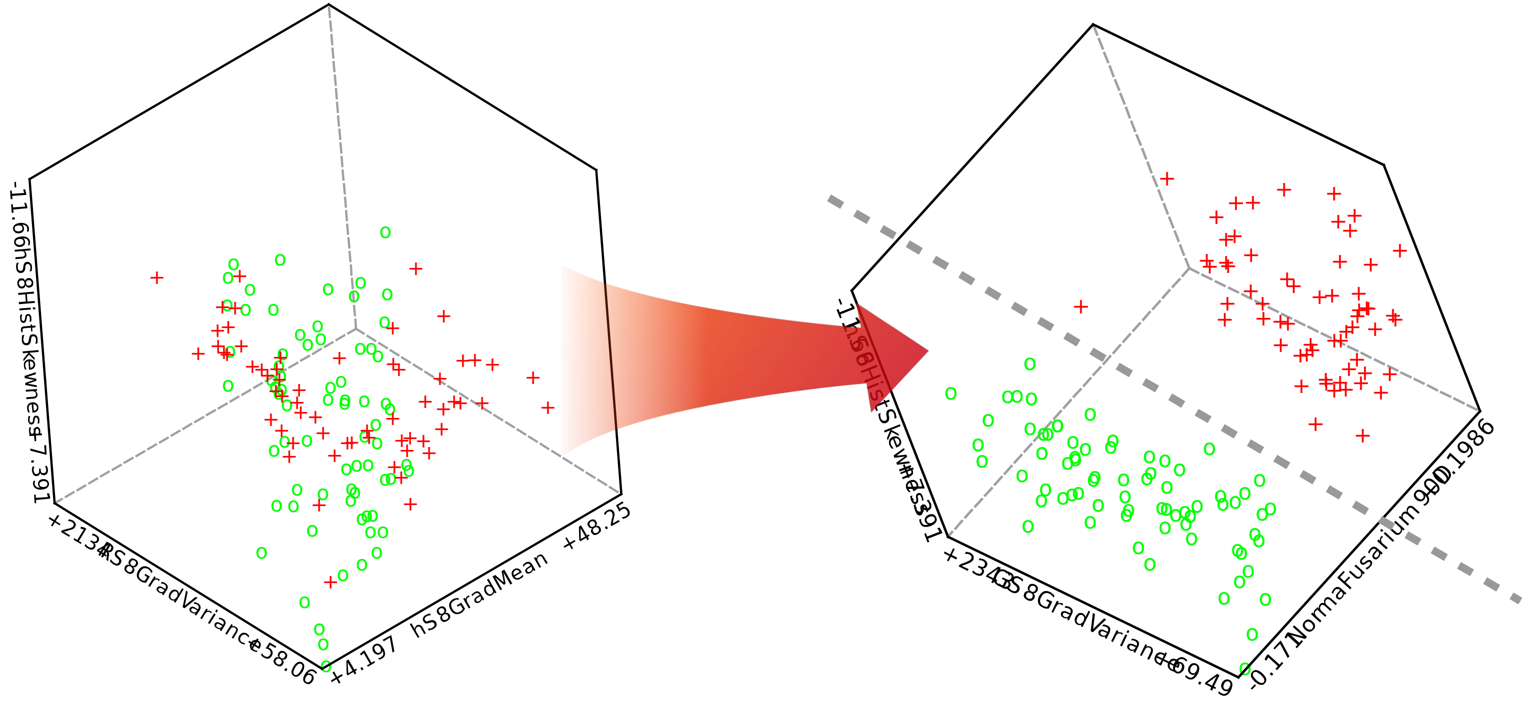
Space of three features
o vectors of fusarium
+ normal kernels

Technical University of Lodz
Institute of Electronics

Machine learning explores algorithms that from example inputs create data-driven prediction models or decisions rules.

Linear discriminant analysis is a method to find a linear combination of features that separates two or more classes of objects.

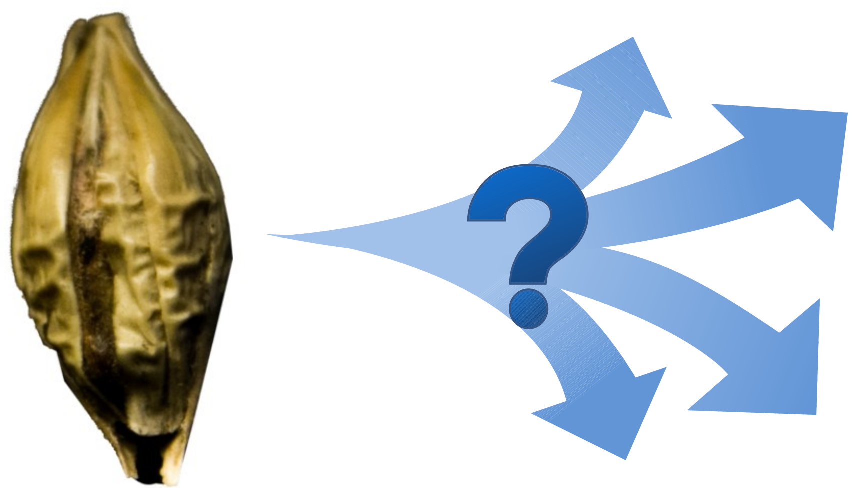
The resulting combination is used as a **linear classifier**, in which a decision boundary is simply a line (2D), plane (3D) or hyperplane.



Data classification

Classification is a process of identifying to which category (class) a new observation (kernel) belongs.

Linear **classifier ensembles** are applied to deal with multi-category identification.



Good
Sprouted
Missing germ
Fungal infections
...

Cross-validation

Cross-validation is a technique for estimating the performance of a classifier. It involves partitioning data (set of feature vectors) into subsets, performing machi-ne learning on one subset and validating the classifier on the other one.

In this study the three-fold cross-validation was applied:

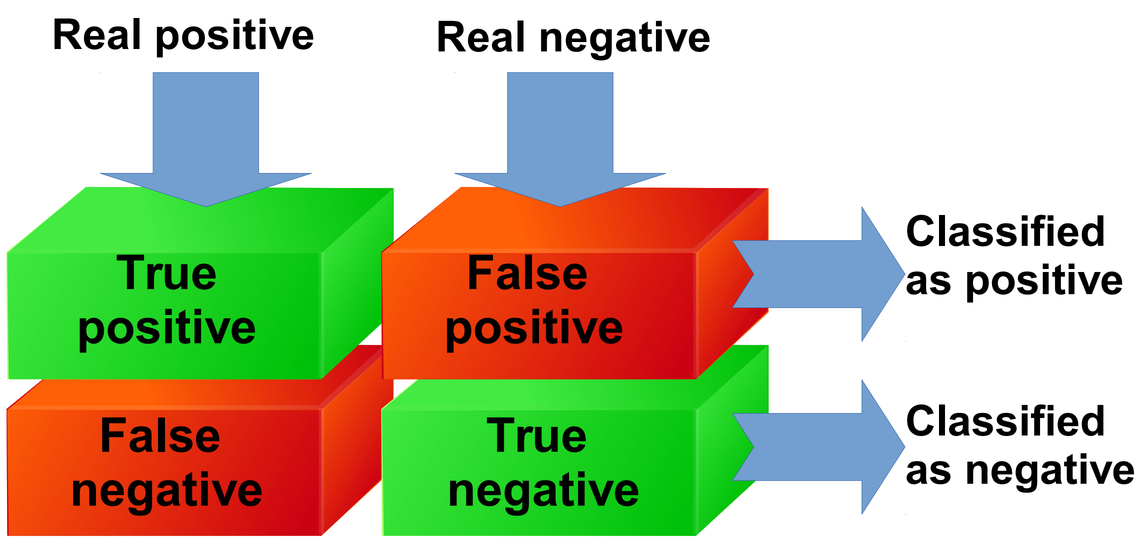
- vector sets were randomly divided into three possibly equinumerous subsets
- two subsets were used for training and remaining one for validation
- the process was repeated for each of the three subsets used for validation

Sensitivity and specificity

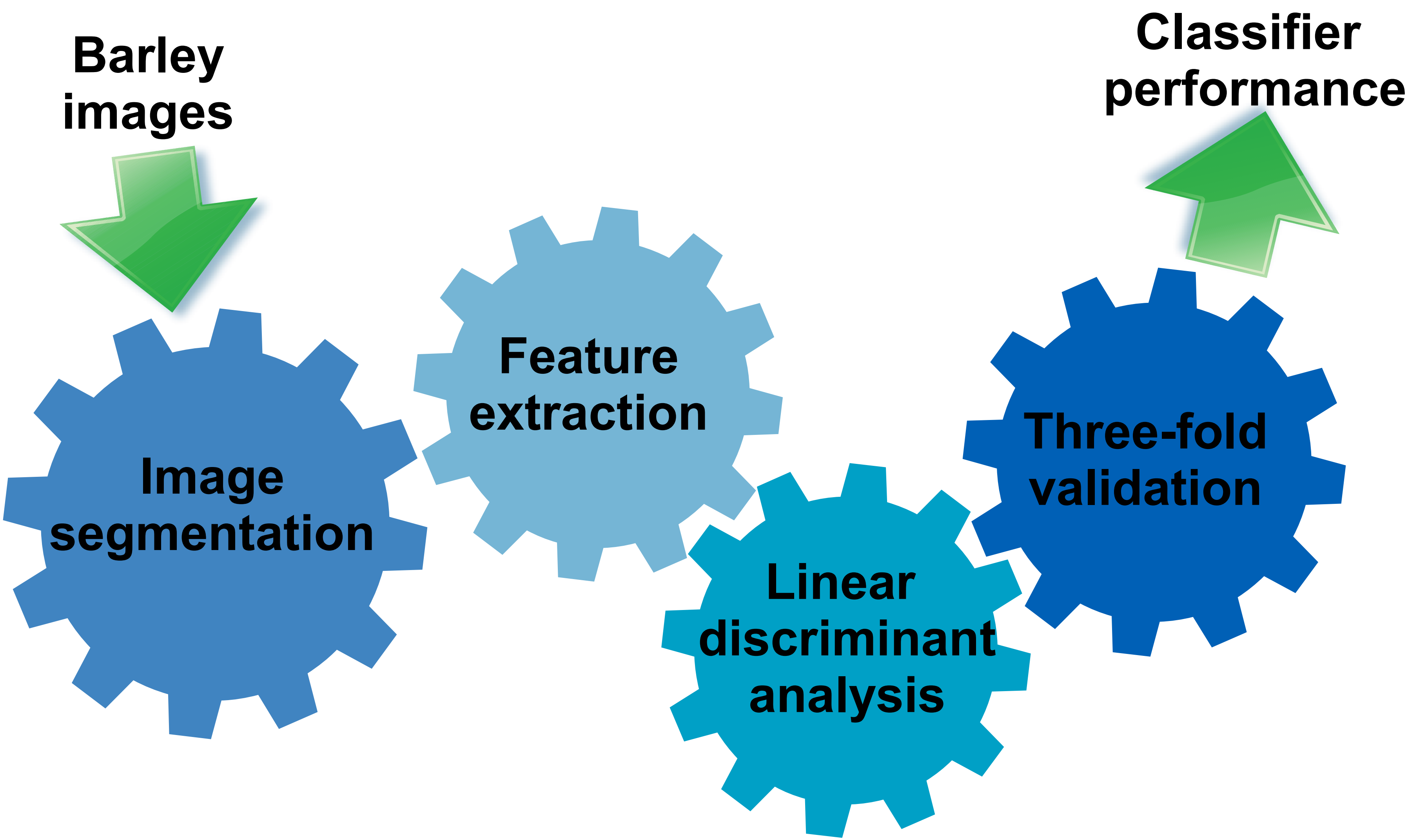
Sensitivity and specificity are statistical measures of the performance of a binary classification test [wikipedia].

$$sensitivity = \frac{number\ of\ true\ positives}{number\ of\ true\ positives + number\ of\ false\ negatives}$$

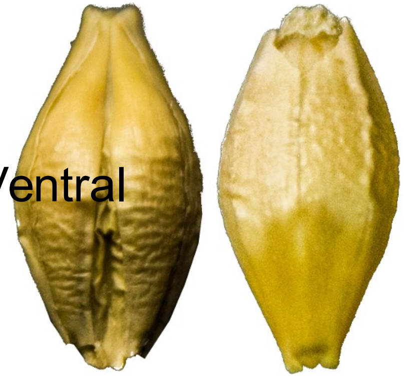
$$specificity = \frac{number\ of\ true\ negatives}{number\ of\ true\ negatives + number\ of\ false\ positives}$$



Methods


Barley images
Image segmentation
Feature extraction
Linear discriminant analysis
Three-fold validation
Classifier performance

Which side is better?



Ventral

Dorsal

Missing germ	Good for malting	Ventral
28	5	Missing germ
2	30	Good for malting

Sensitivity: 0.93
Specificity: 0.86

Missing germ	Good for malting	Dorsal
30	1	Missing germ
0	34	Good for malting

Sensitivity: 1.00
Specificity: 0.97

Linear classifier performance

Missing germ	Good for malting	Ventral
30	1	Missing germ
0	34	Good for malting

Sensitivity: 1.00
Specificity: 0.97

Fusarium	Good for malting	Ventral
40	4	Fusarium
2	31	Good for malting

Sensitivity: 0.95
Specificity: 0.89

Fungal infection	Good for malting	Dorsal
52	6	Fungal infection
2	29	Good for malting

Sensitivity: 0.96
Specificity: 0.83

Classifier ensemble performance

Swollen	Fusarium	Fungal inf.	Mis. germ	Good	Dorsal
18	1	1	3	4	Swollen
0	34	1	1	0	Fusarium
2	0	42	0	8	Fungal inf.
1	0	0	25	0	Mis. germ
3	0	3	0	17	Good
6	7	7	1	6	Unassigned

Swollen	Fusarium	Fungal inf.	Mis. germ	Good	Ventral
17	0	1	0	2	Swollen
0	35	0	1	1	Fusarium
4	1	46	3	1	Fungal inf.
0	0	1	21	2	Mis. germ
5	4	1	0	21	Good
4	2	5	5	8	Unassigned

Conclusions

- Computer vision methods can identify barley kernels suitable or unsuitable for malting.
- Combination of image segmentation, feature extraction and machine learning methods does the job.
- Accuracy of classification (the average of sensitivity and specificity) ranges from 0.88 to 0.98 for linear classifiers and reaches 0.72 for classifier ensembles.
- Analysis of dorsal or ventral side of kernels has impact on classification results.