Article Identifier
DOI: 10.22032/dbt.58855
URN: urn:nbn:de:gbv:ilm1-2023isc-049:9

60th ILMENAU SCIENTIFIC COLLOQUIUM
Technische Universität Ilmenau, 4 – 8 September 2023
URN: urn:nbn:de:gbv:ilm1-2023isc:1

Intelligent Classification and Data Augmentation for High Accuracy AI Applications for Quality Assurance of Mineral Aggregates

Katharina Anding,[1,4], Galina Polte[1], Daniel Garten[2], Elske Linss[3], Gunther Notni[1]

[1] Ilmenau University of Technology, Faculty of Mechanical Engineering, Group of Quality Assurance and Industrial Image Processing, Gustav-Kirchhoff-Platz 2, 98693 Ilmenau, Germany
[2] GFE Schmalkalden e.V., Näherstiller Straße 10, 98574 Schmalkalden, Germany
[3] Materialforschungs- und -prüfanstalt at Bauhaus-University of Weimar, Coudraystraße 9, 99423 Weimar, Germany
[4] Now with Fraunhofer Institute for Digital Media Technology IDMT, Industrial Media Applications, Ehrenbergstraße 31, 98693 Ilmenau, Germany

ABSTRACT

In this work, a method for automatic analysis of natural aggregates using hyperspectral imaging and high-resolution RGB imaging combined with AI algorithms consisting of an intelligent deep-learning-based recognition routine in form of hybrid cascaded recognition routine, and a necessary demonstration setup are demonstrated. Mineral aggregates are an essential raw material for the production of concrete. Petrographic analysis represents an elementary quality assurance measure for the production of high-quality concrete. Petrography is still a manual examination by specially trained experts, and the difficulty of the task lies in a large intra-class variability combined with low inter-class variability. In order to be able to increase the recognition performance, innovative new classification approaches have to be developed. As a solution, this paper presents an innovative cascaded deep-learning-based classification and uses a deep-learning-based data augmentation method to synthetically generate images to optimize the results.

Index Terms **–** petrography, deep learning, cascaded AI, data augmentation

1. INTRODUCTION AND STATE OF THE ART

Mineral aggregates are an essential raw material for production of concrete for urban buildings. They are used as concrete aggregates and must be analyzed according to certain European standards, in Germany also according to relevant federal standards, for example [1] in Saxony-Anhalt. Samples from natural mineral deposits must be taken at regular intervals in the deposits to analyze the mineral components. This petrographic analysis represents an elementary quality assurance measure for the production of high-quality concrete. Deposits that exceed the specified limits for the presence of certain minerals are not suitable for concrete production. In Germany and worldwide, numerous structural damages are caused by using constituents in natural aggregates that are harmful to concrete (for example opal sandstone, flint etc., see Figure 1). Using alkali-silica reactive aggregate, which are harmful to concrete, leads to alkali-silica reaction (ASR). A selection of other critical classes is also shown in Figure 1.

Figure 1: Examples of critical aggregates in concrete production (from left to right: opal sandstone, flint, chalk, coal, pyrite / marcasite, brown iron)

The state of the art for a petrographic analysis of natural aggregates is still a manual examination by specially trained experts (mineralogists or geologists). The manual process of petrographic analysis using a manual examination as well as scratch test and hydrochloric acid swab test for questionable stones. This procedure is time-consuming and error prone and their results depend on the individual experience of the inspector. Automatic routines can only be found for the analysis of a few specific mineral components (for example Calcite and Dolomite) by using the near infrared spectra (NIR) [2].

One publication by Wotruba et al. explains that the number of minerals, which can be automatically analysed in NIR, is limited because not all minerals showing diagnostic absorption features and very dark minerals can also cause problems by absorbing a high amount of the electromagnetic spectra [2]. The analysis of micrographs is another way for automatic classifying of natural aggregates [3]. This procedure is used for the inspection of grinding surface patterns of mineral plates and needs a costly mechanical preprocessing of the mineral surface. But for the classification of natural aggregates as untreated rock material, this type of procedure is much too complicated and therefore not suitable for the given task of automating a petrographic analysis.

Our paper describes an automatic recognition routine for petrographic inspection by using image processing of hyperspectral imaging data and AI algorithms in the form of pre-trained convolutional neural networks in the SWIR spectrum.

In our own previous studies, first we used image processing and machine learning based on color images in the spectrum of the visible wavelength range (VIS) [4], [5], [6] and later we used hyperspectral imaging data and deep learning (DL) [7], [8]. In our previous studies using RGB-color images as data base especially the differentiation of very similar classes and their correct assignment to the superordinated-classes of uncritical and two types of critical classes could not yet be solved sufficiently well. For this reason, further investigations in SWIR were continued in the form of hyperspectral data analysis using DL [7].

Hyperspectral data analysis also known as hyperspectral imaging (HSI) combines spectroscopy and computer vision (CV). HSI allows the analysis of both spatial and spectral properties of objects. A hyperspectral image contains information about the spatial distribution of materials within a sample. Each image pixel represents an individual point on the surface of the sample, which includes a quasi-continuous point spectrum of the material [7].

There existing several techniques for capturing hyperspectral images. One of the popular techniques, which is also used in industrial context, is called a push broom imager. Using this technique, images are captured line by line as the object moves relative to the camera. These lines contain a spatial and a spectral dimension. To obtain a whole image, another spatial dimension is added by stitching these lines together [9].
Overall, a hyperspectral image can be thought of as a stack of grayscale images where each grayscale image represents the spatial reflectance intensity per wavelength of a given sample (Figure 2).

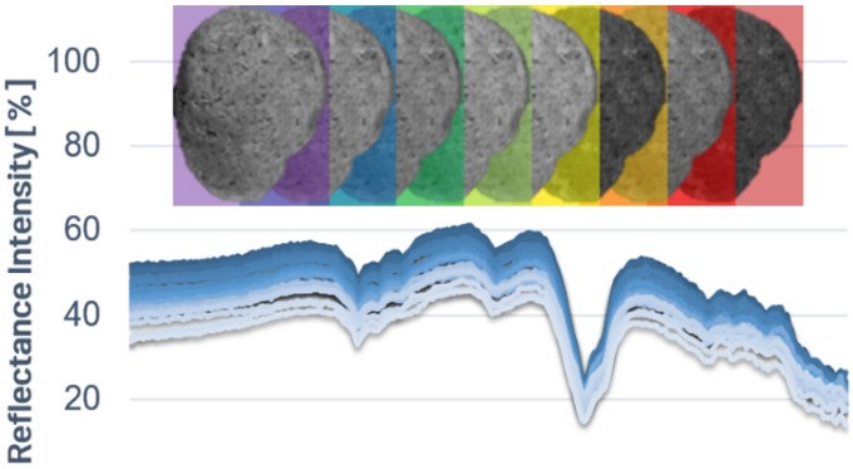


Figure 2: Hyperspectral images of given natural aggregate – Spatial and spectral characteristics [8]

It is important to know, that some of these grayscale images contain redundant information. Therefore, it is important to reduce the dimension of the data using feature extraction and / or feature selection techniques.

Another underlying problem of the given recognition task is that natural aggregates are already present within individual classes as mixed mineral components with specific chemical and spectral properties (a rock particle often contains several mineral components in its structure). Therefore, it is important to develop more complex Artificial Intelligence (AI) based recognition routines to reliably detect the large number and high intra-class variance of potentially occurring classes in a deposit as well as the presence of rock combination within a rock object.
This explains why determining the class of rock particles using near infrared (NIR) and short-wave infrared (SWIR) characteristics is much more complex than separating pure minerals. In the analysis of natural aggregates, the difficulty of the task lies in a large intra-class variability with simultaneously low inter-class variability. In terms, this means that the individual class shows strong variations in appearance and the differences in appearance between the classes are unfortunately often very small. This underlying characteristic of the given recognition task leads to possible confusion between several classes. In order to be able to increase the recognition performance despite the given complexity, innovative new approaches to classification have to be developed. As a solution, this paper presents an innovative AI based cascaded classification that builds on previous research [4], [5], [6] and [7] and combines different classifiers.
Another problem prevalent in recognition of mineral aggregates, the under-representation of rare but reliably recognizable classes, is also addressed in this paper. This problem is of particular importance in the context of the application of powerful deep neural networks due to

the very large amounts of data required by a deep learning network for all classes to be recognized with high confidence.
To solve the main problem of deep learning: the required large training data sets and underrepresented classes, different approaches have been developed from researchers in the past. While Transfer Learning [10] was introduced to solve similar recognition tasks using the same network and training of the fully connected classification layers, dataset augmentation techniques were developed to artificially increase the number of objects present by simulating and varying the objects in a virtual space using simple algorithms. This increase in the total number of objects within a dataset reduces the risk of overfitting when training a deep neural network and reduces the cost of creating a homogeneous dataset [11].
Classical methods of dataset augmentation include cropping, mirroring, scaling, random erasing [12], rotational or translational motion [13]. An innovative approach is the application of Generative Adversarial Networks (so-called GANs). These network architectures learn to simulate images based on a given dataset of real objects and thus generate new artificial images that take into account the specific characteristics of the given real dataset examples. GANs describes a supervised learning algorithm, used for data augmentation [14]. The algorithm consists of a pair of networks, the so-called Generator and Discriminator. Both work against each other while training process, the generator creates an image based on a random noise pattern while the Discriminator is given either the fake images created by the Generator or a real image from the dataset and decides whether the image given is real or fake. This decision is given as feedback to the Generator. Therefore, while the Generator learns to create more realistic images, the Discriminator becomes better in distinguishing between fake and real.
This paper shows the solution of automated rock constituent recognition using intelligent cascaded classification and presents a possibility of recognition rate increase using an intelligent data augmentation method based on Deep Learning. In the paper, a comparison without data augmentation and with data augmentation is shown to highlight the possibilities and limitations.

2. INVESTIGATIONS

2.1 Image acquisition with push broom sensor in the wavelength range of 960 – 2500 nm

The hyperspectral images of various natural aggregates were taken in a selected spectral range on a hyperspectral workbench from the manufacturer HySpex / NEO (HySpex SWIR-384) (Figure 3).

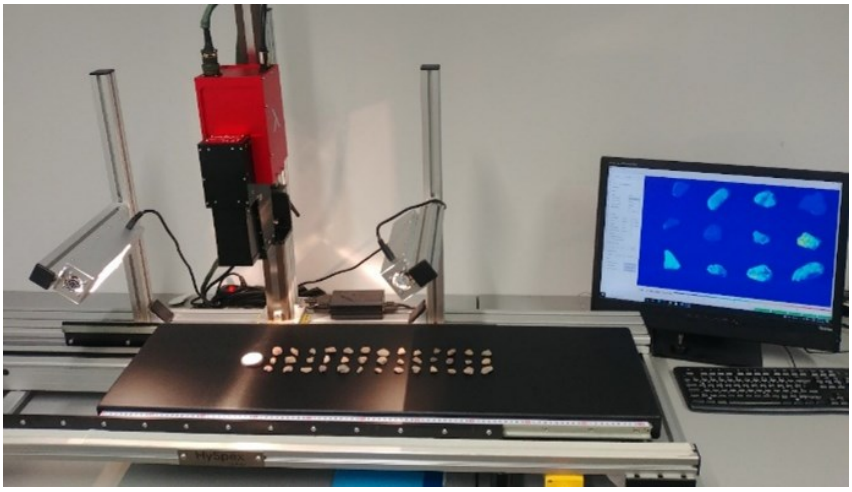


Figure 3: Used push broom sensor HySpex SWIR-384

The used SWIR camera works on the push broom imaging principle and covers a spectral range from 960 nm to 2500 nm. The spectral resolution is about 5.45 nm with 288 image channels (spectral channels), a spatial resolution of 384 pixels per line and a frame rate from up to 400 fps. The illumination in form of two halogen lamps covers a wavelength range from 400 nm to 2500 nm.

2.2 Image acquisition with high-resolution color camera line sensor in VIS

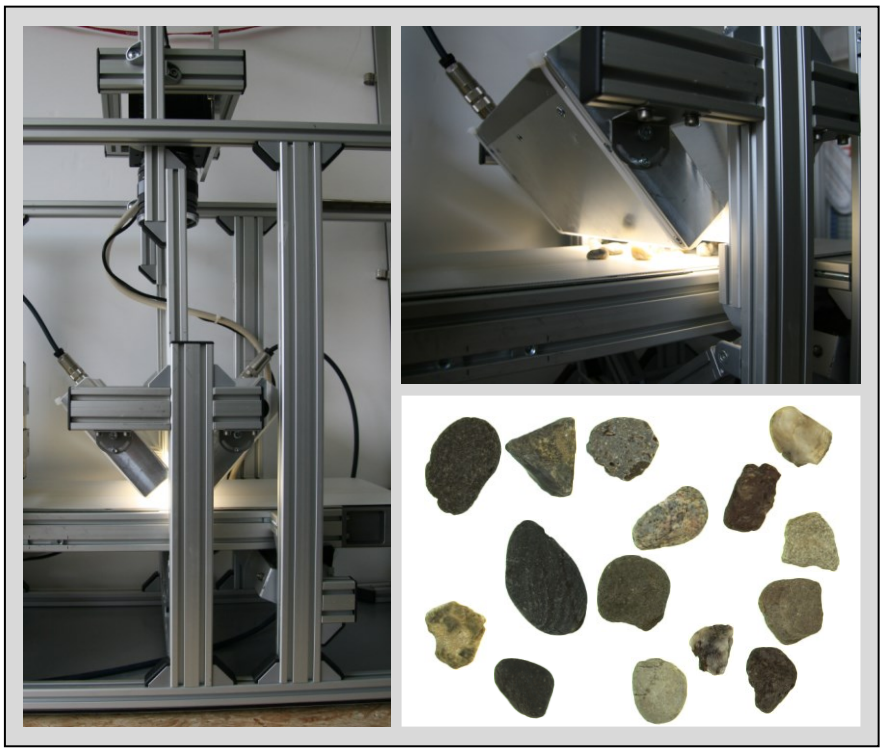


Figure 4: Used RGB color line scan camera

The RGB image partial data sets of the investigation were generated using an experimental setup consisting of a 3-fold color line scan camera with divisor prism as well as a suitable incident light illumination (see Figure 4). The lens used was an achromatically corrected macro lens. An LED line light with a high homogeneity along the line was used as the illumination device. The big advantage of the RGB color camera images in contrast to the HSI images is the significantly higher spatial resolution (2048 pixel per line), which offers an advantage for a supplementary classification based on the object textures, especially for classes that are difficult to separate.

2.3 Acquired datasets

In our research, we used a large HSI data set from our study in [7]. The examples of the different subclasses were collected by our partner at MFPA Weimar, recorded with the HSI Sensor, analyzed and structured. Table 1 gives an overview of the samples of collected natural aggregates in the 3 main classes (super-ordinated classes) and the 21 categories (sub-classes).

Table 1: Overview of the investigated categories of natural aggregates [7]

Category	Aggregates	Name of class
1	Quarz	non-critical aggregates
2	Siliceous slate (black, grey)	
3	Quartzite	
5	Other Palaeozoic sediments (quartzitic, clay, phyllitic schists)	
6	Sandstone (dense) except group 16	
7	Limestone (marlstone), native except group 15**	
8	Limestone (dolomite), Nordic except group 15**	
10	Crystalline (granite, gneiss), Nordic	
11	Flint (dense), all varieties except group 12*	
4	Greywacke	critical aggregates 1 (Water-absorbent, swellable inorganic, alkali-reactive)
9	Rhyolite, andesite, (porphyries, porphyrites), basic volcanic rocks	
12	Chalk-crusted and porous flint (flint)*	
13	Siliceous limestones, siliceous chalk, opaline sandstone	
14	Chalk / Chalk Limes	
15	Light and porous limestone and	
16	Sedimentary rocks with loose grain bonding (e.g. mudstone / siltstone /	
17	Brown coal	critical aggregates 2
18	Charcoal wood, xylitol	
21-1	Other - plastics, rubber, polystyrene	
21-2	Other - organic material (plant parts, soil, ...)	
19	Brown iron encrustations, turf iron ore	
21-3	Other - Goethite	
21-4	Other - Haematite	
21-5	Other - metals	
21-6	Other - building materials (mortar, concrete, bricks,...)	
20	Pyrite, marcasite	
21-7	Other - gypsum	

The super-ordinated class “non-critical aggregates” consists of natural aggregates, which are harmless for production of concrete, “critical aggregates 1” consists of alkali-silica reactive, water-absorbing, and swellable inorganic compounds with the potential for concrete destruction, and “critical aggregates 2” consists of organic, ferrous and sulphur / sulfate compounds with critical potential, too.
In Figure 5 the object numbers in individual categories of the large HSI data set are shown. Figure 6 shows the studied object number in individual categories for the classes category 7 (Cat_07: non-porous limestone) and category 15 (Cat_15: light and porous limestone) of the RGB dataset, which are used for studies of dataset augmentation with GAN.

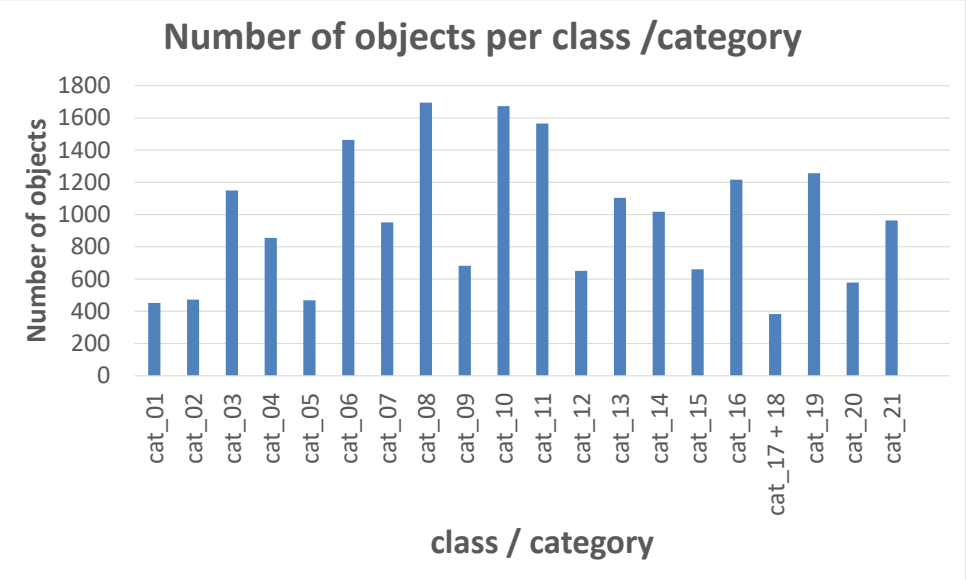


Figure 5: Used HSI dataset with sub-classes / categories

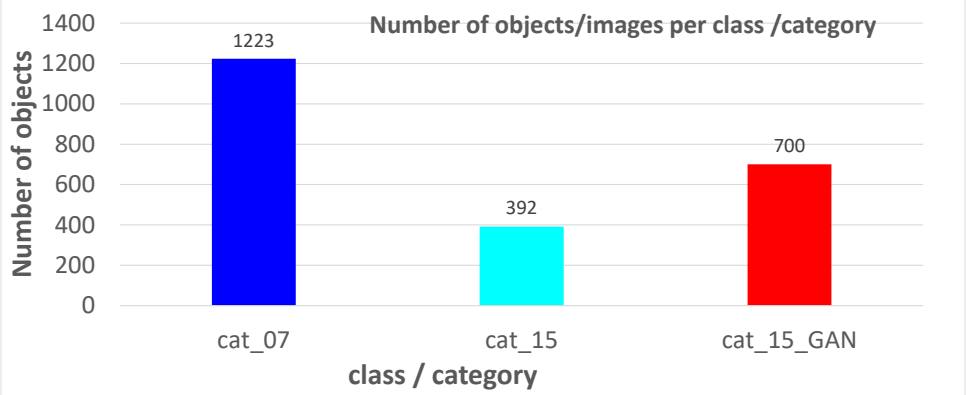


Figure 6: Used RGB dataset with sub-classes / categories (original and GAN)

2.4 Image processing and cascaded AI recognition routine

The image processing and AI recognition routine were developed and validated at TU Ilmenau and the following steps (chapter 2.5 – 2.9) were also done at TU Ilmenau. The first step is the acquisition of HSI Data. The second step is summarized as so-called pre-processing, which consists of: normalisation and baseline correction of the data, definition of an ROI in spatial and spectral dimension, and outlier analysis, a data reduction method and at the end the segmentation process. The pre-processing steps are necessary due to the high dimensionality of the hyperspectral data before applying a deep learning model. After pre-processing, one has to decide between a conventional classifier (e. g. Random Forest) or a Deep Learning classifier (CNN). When using a conventional classifier, the next steps are feature extraction, feature selection, classifier training, testing and validation. When using a DL classifier, on the other hand, feature extraction and selection are already part of the intrinsic DL algorithm, so no prior feature extraction or selection is required.

Unfortunately, the HSI sensors available on the market still have only limited spatial resolution, in contrast to the available high-resolution color camera line sensors and despite their high price. However, since various preliminary investigations of our own revealed the importance of high spatial resolution for distinguishing classes that are phenotypically difficult to separate and can only be distinguished on the basis of their porosity (for example dense and porous sandstone), the chosen approach was to develop a cascaded hybrid detection routine. So we

want to use the benefit of a texture analysis in high-spatial resolution RGB images and the benefit of deep Learning in high-spectral resolution HSI images.
In a cascaded recognition routine, there are two ways of implementation. The first possibility is to align the different sensor signals from the RGB sensor and the HSI sensor beforehand, so that one can perform feature extraction in both sensor data with the same spatial resolution and direct object assignment. Subsequently, the feature vectors extracted in both signals are fused and used in a common classifier (e.g., a support vector machine) as the classification basis. The main drawback of this method, however, is that one forfeits the benefits of extracting high-resolution texture features as a classification basis in favor of a uniform resolution of both sensors, since one can reasonably achieve an approximation of the spatial resolution only to the smaller resolution available (without any real information gain). A second method, which is preferred by the authors of this paper, is to use the output signals in their original resolution in different stages of the recognition routine.
Figure 7 shows possible strategies to realize a cascaded AI-recognition routine using a multi-stage recognition process based on different sensors (multimodality). Thus, in a first step, either a CNN classification can be performed on the high spectral resolution HSI image data after prior application of an LDA or, alternatively, a CNN classification can be performed on the high spatial resolution RGB image data. In a second detection stage there are 3 alternatives:

1) a classification using a conventional classifier (e.g.: Random Forest) based on the point spectra of the HSI data or
2) a classification by conventional classifier (e.g.: Random Forest) based on selected wavelength bands of the HSI data or
3) a classification by conventional classifier (e.g.: Random Forest) based on a previous color/gray value and texture analysis of the RGB data.
   In this way, the respective advantages of the sensors can be optimally exploited in order to reliably distinguish even classes that are difficult to separate.

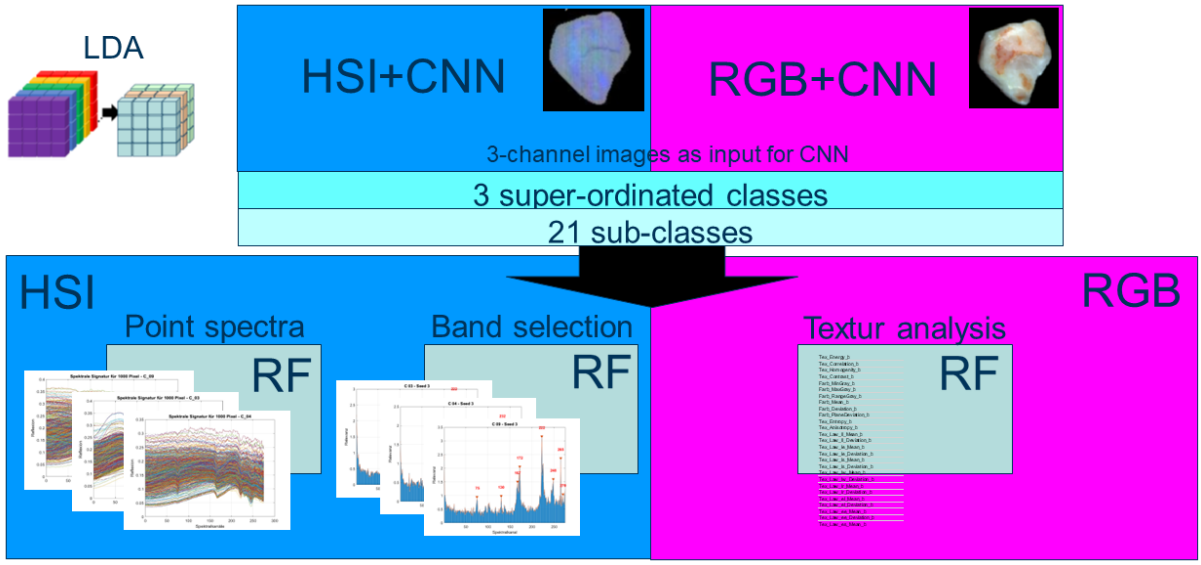


Figure 7: Cascaded AI-recognition routine

For example (see Figure 8), in a first step one can perform a deep-learning classification on the HSI data after applying an LDA, and for objects classified into a class with low confidence level (which are difficult to separate), in a further step we propose a texture feature analysis and classification with a conventional classifier (e.g., an SVM or a Random Forest) on the high spatial resolution color camera data. Considering the final result, the two classifiers then act

similar to an ensemble classifier and a hybrid determined classification result is obtained with a higher confidence level.

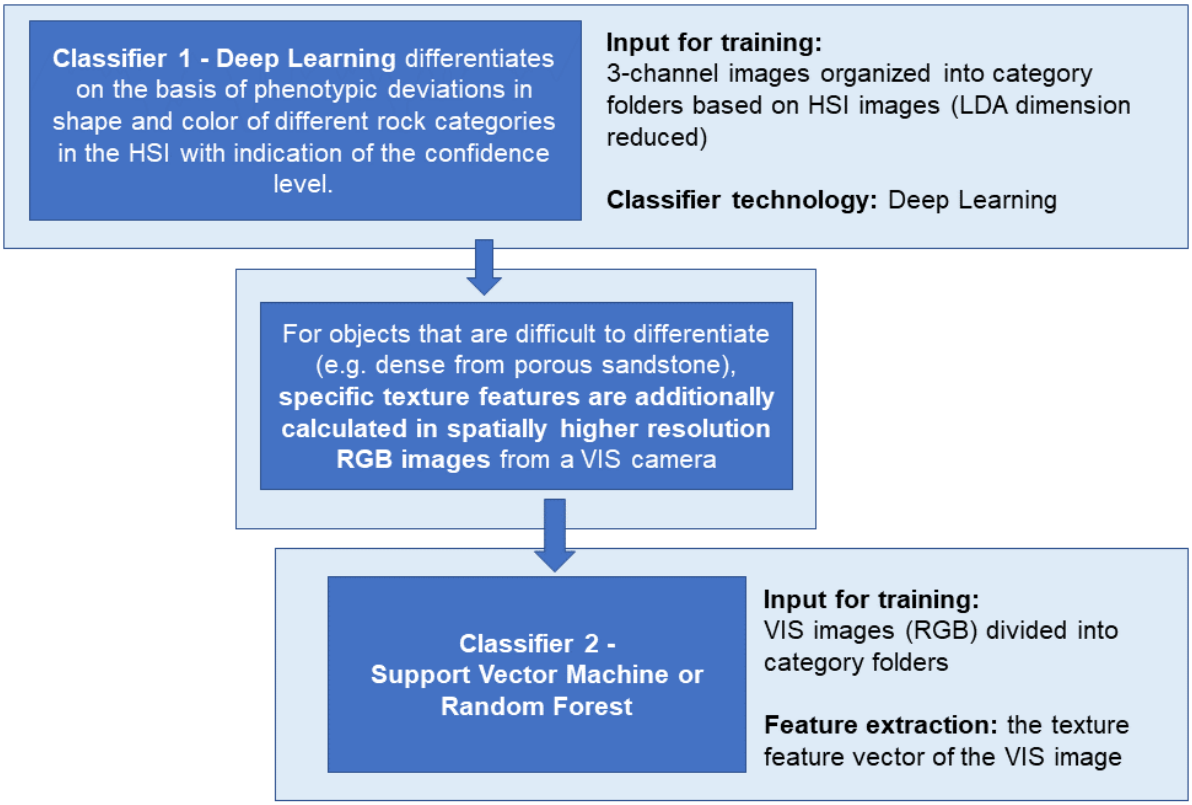


Figure 8: Cascaded hybrid recognition routine on the specific example

2.5 Dimensional reduction of hyperspectral data for the application of Deep Learning

Within the scope of the investigations, the aim was to apply pre-trained CNNs to the acquired hyperspectral images for getting good results. Before the actual application of a CNN for the analysis of the aggregates, it is necessary to process the hyperspectral images with different pre-processing steps and to segment the actual object regions. A segmentation method was implemented using a pixel-based thresholding. The segmented aggregates are then transformed from a higher dimensional space to a three-dimensional space using various methods of multivariate data analysis (dimensionality reduction techniques) (resulting in 3-channel images) to make them useful for classification with pre-trained deep learning models. The dimensionality reduction methods investigated were principal component analysis (PCA), independent component analysis (ICA), and linear discriminant analysis (LDA). For the given application, linear discriminant analysis (LDA) shows the best results.

Figure 9 shows the false-colour representation of natural aggregates after application of the LDA with regard to the superordinated-classes “non-critical aggregates”, “critical aggregates 1” and “critical aggregates 2”.

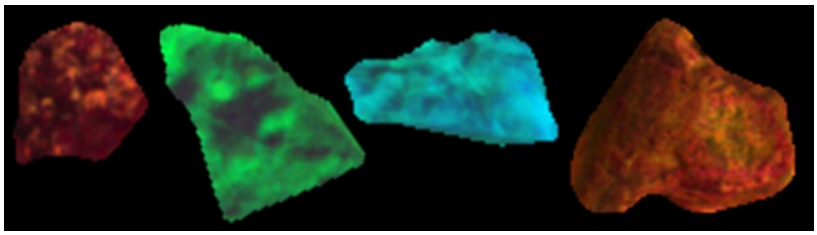


Figure 9: False-color representation of natural aggregates after application of the LDA with regard to the super-ordinated classes

In Figure 10, the results of dimensional reduction using Linear Discriminant Analysis (LDA) are visualized in the form of the outlier-cleaned point spectra after LDA.

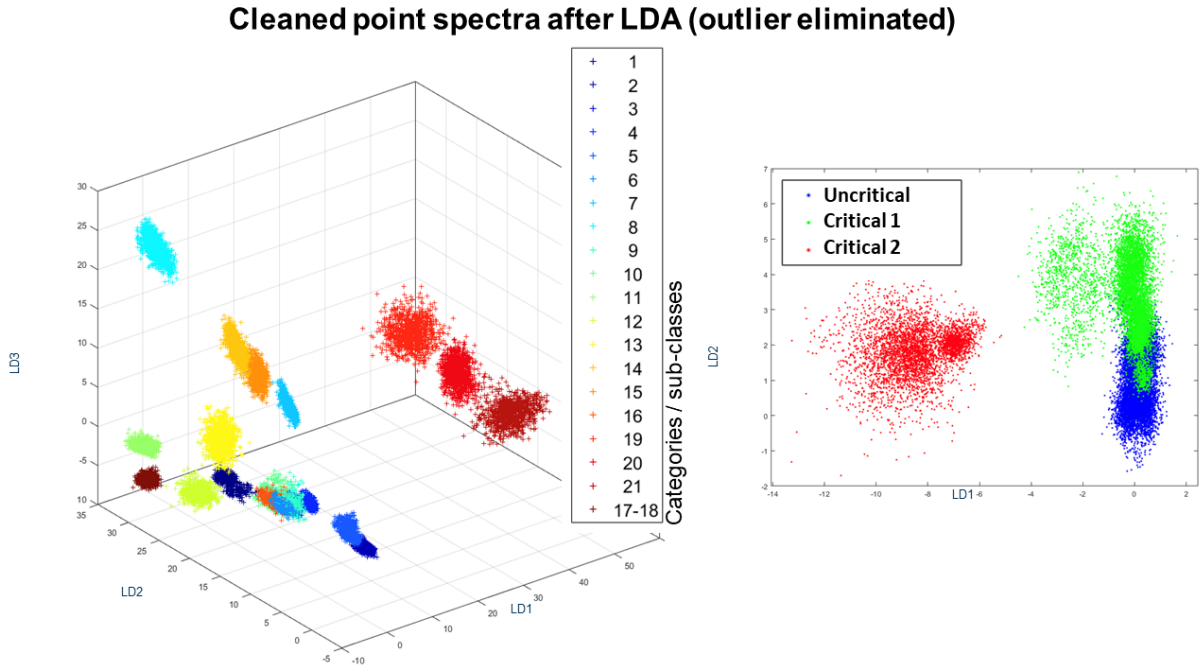


Figure 10: Dimensional reduction using Linear Discriminant Analysis (LDA) visualized in the form of the outlier-cleaned point spectra after LDA [15]

2.6 Training, validation and test of preselected CNN models on dimension-reduced hyperspectral images

For the classification of the dimension-reduced hyperspectral images, different pre-trained deep-learning models in the form of so-called Convolutional Neural Networks (in short: CNNs) were first used (VGG-16, ResNet-50, CNN Halcon-enhanced and MobileNet), which were developed specifically for the classification of 3-channel image data [7], [8]. Thus, using a feature space transformation and dimension reduction by means of LDA and subsequent application of a DL network (ResNet50), a good overall recognition rate of 90.9% on 3 super-ordinated classes ("uncritical", "critical_1" and "critical_2" could already be achieved so far (see Figure 11).

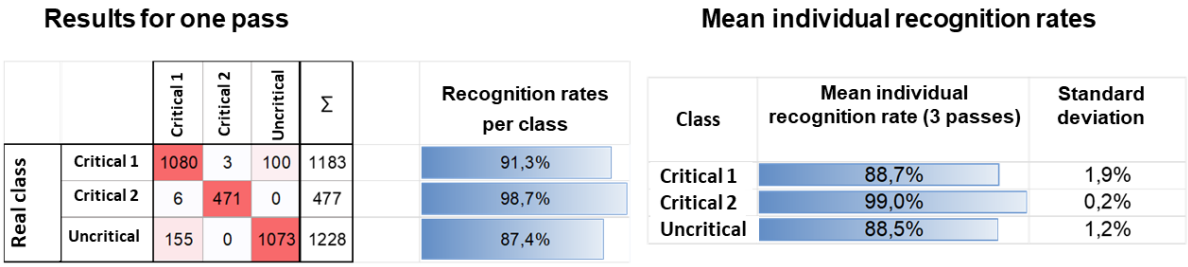


Figure 11: LDA+CNN - Results for the three super-ordinated classes [15]

The goal was to further optimize the recognition performance beyond the 90.9% achieved by using additional classifiers for the objects that were classified with lower recognition confidence in the result of the CNN. The cascading of the recognition routine was aimed at and for this purpose further AI variants were investigated in order to be able to recognize classes that are difficult to separate reliably using an ensemble of classifiers. The cascaded recognition routine was successfully developed as program scripts. In the final data set, which is currently still to be recorded with the final demonstrator (described in chapter 2.10), the investigations carried out are to be repeated accordingly and the developed algorithms of the cascaded AI are to be adapted and optimized to the final sensor signals.

2.7 GAN-based Data Augmentation

To train deep learning networks, a large amount of data is needed to avoid overfitting and to allow the network to have a good generalization capability. Therefore, in addition to the uncritical objects, a large amount of critical class objects is especially needed. For recognition tasks, obtaining a balanced, sufficiently large data set for training presents a particular problem, especially with respect to rare (underrepresented) classes. The collection of underrepresented class examples is costly and time-consuming. However, cost reduction and optimization of recognition performance can be achieved, e.g., by the complementary application of artificial intelligence in the form of Generative Adversarial Networks (GANs) by using them to artificially generate images of underrepresented classes. After training, the GAN network architectures learn to simulate artificial images based on a given data set.
Therefore, the goal was to investigate intelligent synthetic data augmentation to optimize evaluation algorithms in the form of optimized AI models.
In our studies, we used a GAN to simulate synthetic images of the class cat_15 (light and porous limestone), which previously had a poor individual recognition rate due to a high confusion rate with the class cat_7 (non-porous limestone). The basis of the investigations is the RGB data set with high spatial resolution. We used StyleGAN3 [16] as official PyTorch implementation with an image size of the GAN of 256x256 and the graphic card NVIDIA GeForce RTX 2080 Ti with 11 GB. We started with a pre-trained model LHQ-256 [17] and needed 21 days and 21 hours for getting the results in form of a finally trained GAN-Model (with Kimg 12640), which is useful for an intelligent data augmentation of cat_15. The process of investigation of the influence of data augmentation is explained below:

1) In the test always the same original images were used to give a comparability of the recognition performance! The test images of the classes were already removed at the beginning as independent test partition for the comparative tests, the test images were also not used for training the GAN!
2) GAN (StyleGAN3 [16] - official PyTorch implementation) was trained on 300 original objects of cat_15, which were later also used for training and validation of the classifier to determine the recognition performance. The GAN was trained up to 12640 Kimg.

Using the learned GAN model, 700 GAN objects/images were synthetically generated for cat_15.

3) Next, a CNN classifier was used (CNN of the Halcon software library [18]: pretrained_dl_classifier_enhanced and pretrained_dl_classifier_resnet50) to classify on the original images of the training and validation partition or on the original images of the training and validation partition plus the GAN images and compare the recognition performance.

In Figure 12 different image examples as result of the data augmentation process are shown for sub-class 15 (on the left side the original images and on the right side the GAN-simulated images of cat_15). Already with the eye one recognizes the great similarity between original and GAN-simulated images and thus a qualitatively very well assessable authenticity of the artificially generated images both in terms of color and shape, as well as texture. Especially the difficulty of the simulation of an extremely realistic texture is to be emphasized here, which however succeeds extraordinarily well with the used GAN.

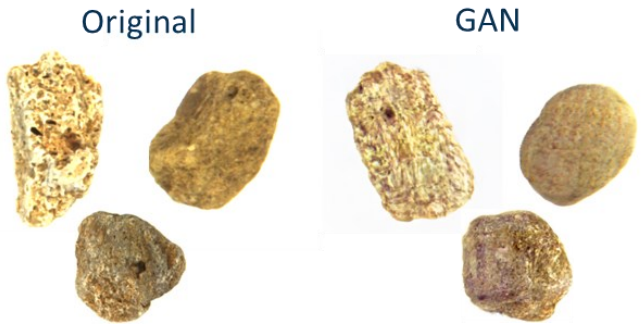


Figure 12: Examples of original and GAN-produced images of category 15 (light and porous limestone)

2.8 Statistical analysis of the quality of data augmentation

In chapter 2.7 the quality of the synthetically, by means of GAN generated images was qualitatively evaluated by eye under evaluation of the gray value, color and texture similarity to original images of the class Cat_15. In this chapter, an approach is made to quantify this qualitative assessment by means of statistical evaluation and to make it measurable. For this purpose, first 123 features of gray value and texture of the two original classes cat_7 and cat_15 (RGB data set) were calculated with Halcon [18] (because the two, visually similar classes differ more in color and texture than in form). The features were extracted on the basis of the HSI color space (components: hue (H), saturation (S), intensity (I) after a transformation from RGB to HSI color space. For example the feature “area_center_gray” was calculated in the three different HSI-channels (H, S, I). This function computes the area and center of gravity of a region in a gray value image and the results are: AreaH, AreaS and AreaI as gray value volumes of the region in channel H, S and I and the coordinates with Row and Column of the gray value center of gravity for each HSI-channel.

Subsequently, the discriminant ability of the features was evaluated using the InfoGain ranker (feature selection method) procedure of the WEKA [19] software library. The ranking result using the 12 best features is shown in Table 2. All 12 best features are results of the gray value features “area_center_gray”, “elliptic_axis_gray” or the texture feature of Laws [20].

Table 2: 12 best features used InfoGain Ranker

Info Gain Value	number in feature vector	Feature name	Halcon function
0.8652	123	rr	texture_laws
0.8146	122	rw	texture_laws
0.8146	2	AreaH	area_center_gray
0.8097	16	RaS	elliptical_axis_gray
0.7999	118	wr	texture_laws
0.7965	22	RaI	elliptical_axis_gray
0.7958	19	RaH	elliptical_axis_gray
0.7865	61	PerimeterS	fuzzy_perimeter
0.7293	23	RbI	elliptical_axis_gray
0.7226	17	RbS	elliptical_axis_gray
0.722	113	sr	texture_laws
0.7187	20	RbH	elliptical_axis_gray

In the next step, the feature distributions of the 12 features with the highest discriminant ability between both classes were examined more closely in their feature distributions under representation of their box-whisker plots. This type of presentation allows a quick overview of the median value, the skewness of the distribution underlying the data as well as occurring outliers. The features sorted in order of discriminant ability using InfoGain are shown by their box-whisker plot in Figure 13. The box plots of the original classes (without data augmentation) are labeled cat_07 and cat_15, the box plot of the data augmented class is labeled cat_15_GAN.

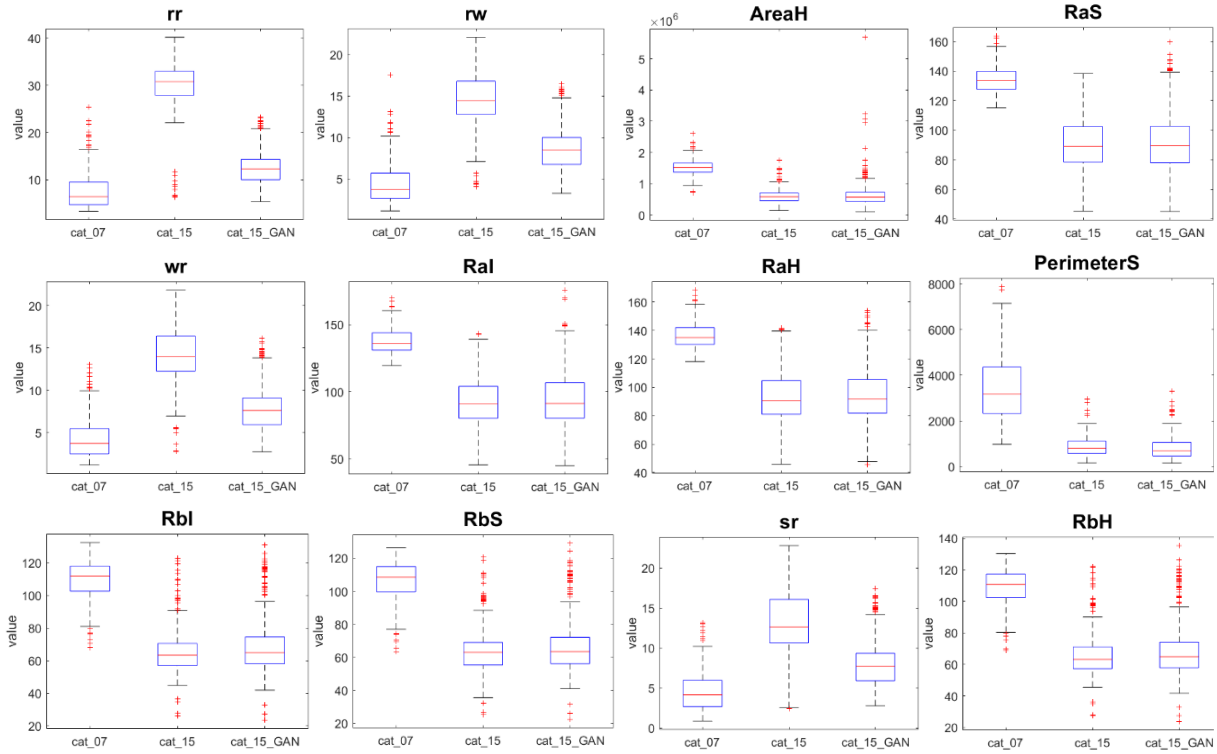


Figure 13:Box-Whisker-Plots of the best 12 features with InfoGain-Ranker

It is notable that for the grey value features AreaH, RaS, RaI, RaH, PerimeterS, RbI, RbS and RbH, there is consistently very high agreement in the box-whisker plots between the original

and data-augmented class of cat_15. Here, both the median values and the position of the boxes relative to each other differ only minimally (nearly the same median and skewness).
In contrast, however, their box-whisker plots differ strongly from the other class cat_07, indicating good separability. For the texture features by Laws (rr, rw, wr and sr), on the other hand, the box-whisker plots of the distribution of original and data-augmented class cat_15 differ strongly. This indicates a not yet optimal simulation of the natural texture of class cat_15 using GAN.
Looking at the outliers, it is noticeable that there are more statistical outliers for the data-augmented class than for the non-augmented class for individual grey value features (e.g.: RaS, RaI or RaH), which speaks for a certain uncertainty of the reproduction of the natural variance of these features. However, for texture features, without exception, there are significantly more outliers in the data augmented to the non-augmented class, again confirming that gray value features can be significantly better replicated/simulated with the GAN than texture features.

Furthermore, the class-specific feature distributions were visualized in histogram representation in the software WEKA for the selected features with high discriminant power. Figure 14 shows the best 12 features in histogram form (cat_07 colored in blue, cat_15 colored in cyan and cat_15_GAN colored in red). In an optimal simulation of the data-augmented images, the two distributions for cat_15 and cat_15_GAN should be almost on top of each other and also clearly different from the distribution of cat_07. This is given for the gray value features. For the texture features, it can also be seen here that the distributions of the data-augmented and non-data-augmented class of category 15 differ significantly with respect to position and amplitude, which speaks for a non-optimal reproduction of the natural texture using GAN. Since the distribution of the texture features of the GAN-augmented class cat_15_GAN even partly shows more similarity to the other class cat_07 than to the original class cat_15, it becomes quite clear here that a reliable distinction of the classes only using the texture features is not given here.

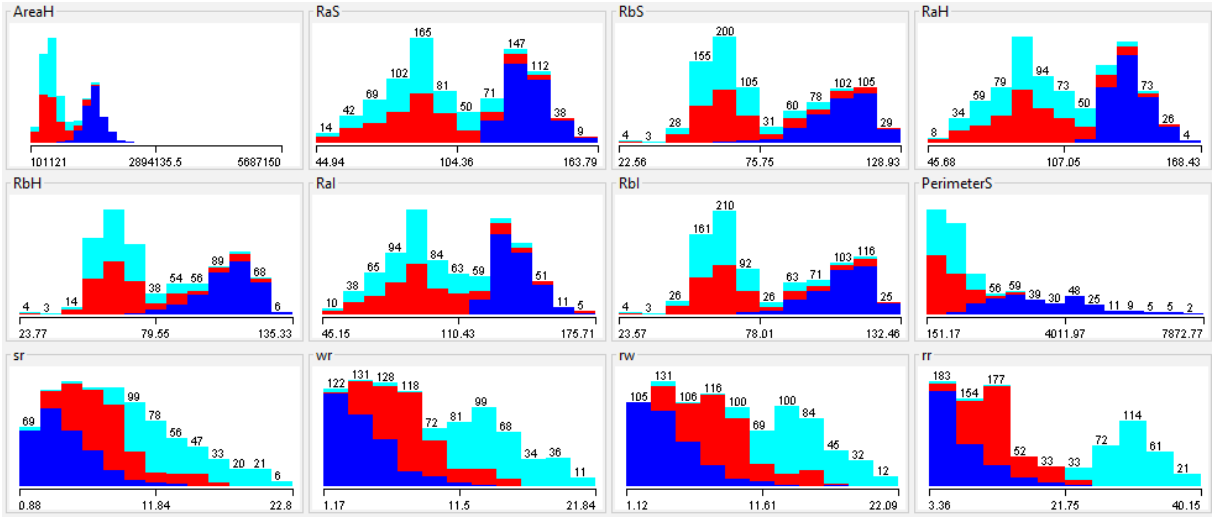


Figure 14: Feature distributions of the best 12 features with InfoGain-Ranker (cat_07 in blue, cat_15 in cyan and cat_15_GAN in red)

2.9 Performance increase of data augmentation

In this chapter the performance increase of the intelligent data augmentation with GAN under the comparative examination on the example of the recognition rates of two different CNNs on a database with and without GAN data augmentation is shown. In the first investigation (see Table 3), the Halcon CNN "pretrained_dl_classifier_enhanced" was used on the non-augmented partial data set, consisting of 1000 original RGB images of the class cat_07 as well as 300 original RGB images of the class cat_15 (without data augmentation) for training and

validation and 223 original RGB images of cat_07 and 92 original RGB images of cat_15 for the test. Furthermore, the same classifier was used in parallel on the GAN-augmented partial dataset, consisting of 1000 original RGB images of the class cat_07 as well as 300 original RGB images plus 700 GAN-augmented RGB images of cat_15 for training and validation as well as 223 original RGB images of cat_07 and 92 original RGB images of cat_15 for testing, were used. The results are shown in Table 3. The classifier "pretrained_dl_classifier_enhanced" showed an absolute recognition performance increase of 0.2% for cat_07 (not significant) and a significant absolute increase of 11.1% for category 15 using data augmentation with GAN.

Table 3: Results with and without data augmentation using Halcon CNN (pretrained_dl_classifier_enhanced)

Nr.	Data Augmentation (DA)	Number of images				Mean individual recognition rate [%]		Standard deviation [%]		Absolute increase in mean individual recognition rate [%]	
		training/validation		test							
		cat_07	cat_15	cat_07	cat_15	cat_07	cat_15	cat_07	cat_15	cat_07	cat_15
1	without_DA (original images)	1000	300	223	92	98.9	65.9	0.8	2.4	-	-
2	GAN DA (original and synthetic images)	1000	1000 (300 original+700 GAN)	223	92	99.1	77.0	0.5	3.3	**0.2**	**11.1**

In the second study, the Halcon CNN "pretrained_dl_classifier_resnet50" was also used on the above-described partial data sets non-augmented and GAN-augmented. The results are shown in Table 4. The classifier "pretrained_dl_classifier_resnet50" showed an even higher absolute recognition performance increase of 0.6% for cat_07 (not significant) and a significant absolute increase of 18.4% for category 15 using data augmentation with GAN.
In both cases, data augmentation with GAN significantly improves the underrepresented class, which speaks for a clear optimization of the CNN model on the basis of balanced classes.

Table 4: Results with and without data augmentation using Halcon CNN (pretrained_dl_classifier_resnet50)

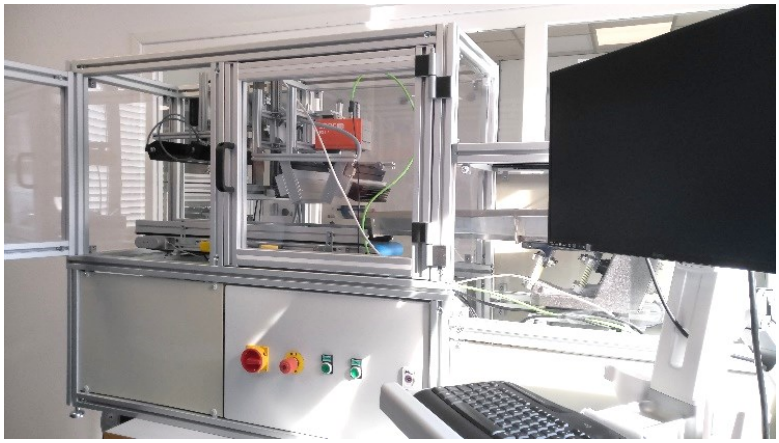
Nr.	Data Augmentation (DA)	Number of images				Mean individual recognition rate [%]		Standard deviation [%]		Absolute increase in mean individual recognition rate [%]	
		training/validation		test							
		cat_07	cat_15	cat_07	cat_15	cat_07	cat_15	cat_07	cat_15	cat_07	cat_15
1	without_DA (original images)	1000	300	223	92	99.0	72.0	0.5	5.4	-	-
2	GAN DA (original and synthetic images)	1000	1000 (300 original+700 GAN)	223	92	99.6	90.4	0.3	5.5	**0.6**	**18.4**

In summary, it can be said that an application of these new techniques for intelligent data augmentation can be applied in industrial practice and can be especially for unbalanced datasets with underrepresented, rarely occurring classes lead to significant increases in recognition rates.

2.10 Demonstrator

In parallel to the software development at the TU, a new demonstrator was built at the project partner GFE - Präzisionstechnik Schmalkalden GmbH, which was realized with optimized components on the basis of the previous test results. Within the scope of the project HyPetro, a demonstrator was designed and set up at GFE GmbH for rock classification as a random sample tester (Figure 15). For this purpose, it was necessary to design a suitable mechanical separation for the strongly heterogeneous samples for individual batches of 5 - 10 kg. Special attention must be paid to the strongly varying densities of the individual sample components, as well as the strong abrasiveness and, in some cases, high porosity of the materials. Furthermore, it must be ensured that the sample components are presented to the VIS color camera and SWIR-HSI camera for data acquisition in a separated and, in particular, overlap-free manner. This will

ensure error-free and stable image acquisition for the generation of learning data sets for the development of the machine learning methods or for the training of the classifier models.
In the final demonstrator, a different HSI camera than the one used in the HSI preliminary investigations (the SPECIM FX17 hyperspectral sensor {see Figure 16} with a wavelength range of 900 to 1700 nm) is used due to lower cost and higher spatial resolution.
As shown in Figure 15, the demonstrator consists of feeding and separation device, conveyor belt, VIS-RGB camera, SWIR-HSI camera and sample discharge as well as the device software for controlling all components.



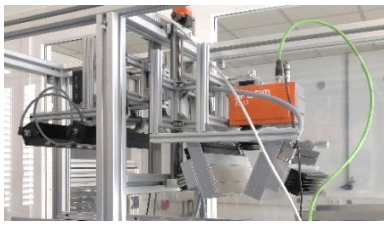


Figure 15: Demonstrator for the detection of rocks in the size range 8-32 mm using VIS-RGB camera and SWIR-HSI camera at GFE - Präzisionstechnik Schmalkalden GmbH

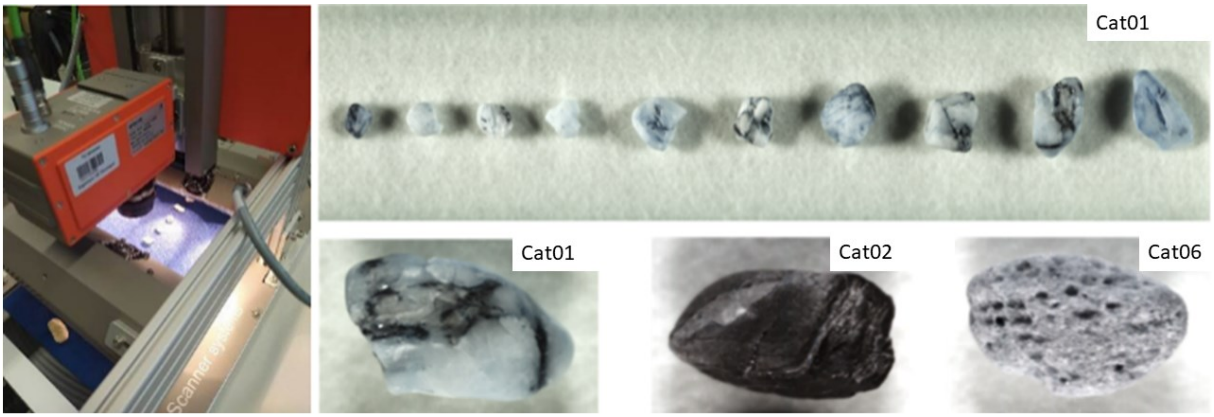


Figure 16: SPECIM FX17 and display of 3-channel image cut-outs of different rock categories on the basis of HSI data

The integration of the classification software developed by the TU Ilmenau is to be carried out in the scope of the project.

3. CONCLUSION AND FUTURE WORK

In this work, a method for the automatic analysis of natural aggregates using hyperspectral imaging and high spatial resolution RGB-Imaging in combination with AI algorithms in the form of an intelligent DL-based recognition routine, a hybrid cascaded recognition routine and a necessary demonstrator setup could be demonstrated. So far, a good overall recognition rate of 90.9 % on 3 superordinated classes ("uncritical", "critical_1" and "critical_2") could already be achieved using a feature space transformation and dimension reduction by means of LDA and subsequent application of a DL network (ResNet50). Furthermore, investigations for an application specific cascaded recognition routine by using complementary classifiers for the objects classified with lower recognition confidence in the result of the CNN, were performed

and in their result the use of a classifier ensemble was elaborated. As soon as the final data set acquired with the final demonstrator is available, there is a need for further optimization within the process chain for the cascaded AI recognition routine, concerning the image acquisition technique, the image pre-processing and segmentation as well as the dimension reduction the learning rate and the different steps of the cascaded routine.
The paper presents different possibilities of cascading the recognition routine and explains their basic steps to be performed. Thus, it gives recommendations for dealing with difficult recognition tasks using multimodal data.
In addition, the paper shows possibilities for dealing with underrepresented classes, which are a widespread problem. A possible performance enhancement of pre-trained deep learning models is shown for synthetic image dataset augmentation using Generative Adversarial Networks (GANs) based on selected investigations. Furthermore, an important aspect of the qualitative and quantitative evaluation possibility of the authenticity of artificially simulated or generated images is investigated and possible evaluation methods are shown.
Further investigations are planned on optimal proportions of synthetically generated images for training in the context of the amount of real recorded images and in dependence on the complexity of the recognition task as well as on the homogeneity of the occurring classes. Further investigations on the evaluability of the image quality of GAN images are also planned. An intelligent synthetic data augmentation of underrepresented object classes will ultimately enable an effective and resource-saving use of the few available objects/instances and thus a more sparing use of resources and time.

4. ACKNOWLEDGMENT

The underlying research and development project "Automated real-time hyperspectral imaging and analysis for the detection of concrete-damaging constituents in aggregates - HyPetro" is funded by the Federal Ministry of Education and Research (BMBF) as part of the funding programme "Computer-Aided Photonics - Holistic system solutions from photonic processes and digital information processing" and supported by the project management organisation VDI Technologiezentrum GmbH.
The specific AI investigations for data augmentation using GAN based on this were performed as part of the ThurAI project, which is funded by the Free State of Thuringia and co-financed by the European Union under the European Regional Development Fund (ERDF).
The computation of the dataset augmentation was shifted from the graphic card to a cluster in further investigations. Special thanks therefore also go to the ProKI network (https://www.proki-ilmenau.de/), in which the TU Ilmenau supports companies in the manufacturing sector in particular with its know-how in the field of artificial intelligence.
We thank all funding sources for their financial support of this work. The funders made it possible to carry out the research on the above-mentioned topics in the first place through their financial support. The responsibility for the research content lies with the authors.

REFERENCES

[1] ZTV-StB LBB LSA 09/10 – “Additionally technical contract conditions and guidance for road construction for the business area State Office for Building Construction of Saxony-Anhalt”, 2009.
[2] H. Wotruba, M. R. Robben, and D. Balthasar, “Near-Infrared Sensor-Based Sorting in the Minerals Industry”, Conference in Minerals Engineering, Lulea, Schweden, 2009.
[3] L. Lepistö, Colour and Texture Based Classification of Rock Images Using Classifier Combinations, doctoral dissertation, Tampere University of Technology, 2006.

[4] K. Anding, D. Garten, G. Linß, G. Pieper, and E. Linß: "Classification of mineral construction raw materials from unconsolidated mineral deposits by means of image processing and machine learning" (Klassifikation mineralischer Baurohstoffe aus Lockergesteinslagerstätten mittels Bildverarbeitung und maschinellem Lernen). 16th Colour-Workshop of the German Color Group, 2010.
[5] G. Linß, K. Anding, D. Garten, A. Göpfert, E. Reetz, and M. Rückwardt, "Automatic Petrographic Inspection by Using Image Processing and Machine Learning", XX IMEKO World Congress, Metrology for Green Growth, Busan, Rep. of Korea, September 2012.
[6] K. Anding, L. Haar, G. Polte, J. Walz, and G. Notni, "Comparison of the performance of innovative deep learning and classical methods of machine learning to solve industrial recognition tasks", In: IMEKO TC1 – TC2 International Symposium on Photonics and Education in Measurement Science 2019, 17.09.-19.09.2019.
[7] K. Anding, G. Polte, P. Hunhold, E. Linss, D. Garten, L. Wunsch, and G. Notni, "CNN-Classification Of Natural Aggregates Used In Concrete Production Based On Their Hyperspectral Characteristics". IMEKO 2022, Porto (Portugal), 30.08. – 01.09.2022.
[8] E. Linß, K. Anding, P. Hunhold, G. Polte, D. Garten, S. Weisheit, and J. Walz, "Automatisierte Analyse natürlicher Gesteinskörnungen mit Betonschädigungspotenzial anhand hyperspektraler Informationen", ibausil Tagung 2023
[9] M. Sackewitz, "Leitfaden zur hyperspektralen Bildverarbeitung", Fraunhofer Verlag, Stuttgart, Germany, 2019.
[10] S.J. Pan, and Q. Yang, "A Survey on Transfer Learning", IEEE Transactions on Knowledge and Data Engineering, Vol. 22(10): p. 1345-1359, 2010.
[11] M. Cogswell, F. Ahmed, R.B. Grishick, C.L. Zitnick, and D. Batra, "Reducing Overfitting in Deep Networks by Decorrelation Representations", ICLR 2016, 2016.
[12] Z. Zhong, L. Zheng, G. Kang, S. Li and Y. Yang, „Random Erasing Data Augmentation", Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34(07), p. 13001-13008, 2020.
[13] S. Yang, W. Xiao, M. Zhang, S. Gou, J. Zhao, and F. Shen, "Image Data Augmentation for Deep Learning: A Survey", arXiv: 2204.08610.
[14] I.J. Goodfellow, et al., "Generative adversarial nets", in Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, MIT Press: Montreal, Canada. p. 2672–2680, 2014.
[15] S. Endreß, Entwicklung einer intelligenten Erkennungsroutine für Hyperspektralbilder unter Verwendung von klassischen, auf Bildmerkmalen trainierten Machine-Learning-Algorithmen. Masterarbeit, TU Ilmenau, 2022
[16] T. Karras, M. Aittala, S. Laine, E. Härkönen, J. Hellsten, J. Lehtinen, and T. Aila, "Alias-Free Generative Adversarial Networks" (StyleGAN3).
[17] J. Pinkney, Model LHQ-256, https://www.justinpinkney.com/
[18] Halcon, Reference Manual - Halcon 23.05, MVTec Software GmbH, www.mvtec.com, 2023, https://www.mvtec.com/products/halcon/work-with-halcon/documentation, 2023.
[19] E. Frank, M.A. Hall, and Ian H. Witten, The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
[20] K. Laws, Rapid texture identification. In SPIE Vol. 238 Image Processing for Missile Guidance, p. 376-380, 1980.

CONTACTS

PD Dr.-Ing. habil. K. Anding email: katharina.anding@idmt.fraunhofer.de
Dr. Galina Polte email: galina.polte@tu-ilmenau.de

