Near-infrared Spectroscopy and Image Classification of Refuse Derived Fuels to Increase Cement Production Quality

Jonas Fischer[1] ✉[0000−0002−9258−4298], Luca Fehler[2], Kevin Treiber[2], and Viktor Scherer[1]

[1] Ruhr University Bochum, Chair of Energy Plant Technology, 44791 Bochum, Germany {fischer,scherer}@leat.rub.de
[2] VDZ Technology gGmbH, 40476 Düsseldorf, Germany
{luca.fehler,kevin.treiber}@vdz-online.de

Abstract. Refuse derived fuels (RDF), produced from municipal and industrial waste, provide an alternative to fossil fuels like coal or lignite in the cement production, thereby reducing the significant CO_2 emissions typically associated with cement production. The composition of RDF is often unknown, which limits the substitution rate, since otherwise the risk of impacting cement quality would increase. In this contribution, both near-infrared spectroscopy (NIRS) and RGB images were used to analyze RDF in an at-line measurement on a conveyor belt setup. The goal was to classify individual RDF particles in one of six fractions (paper, foils, 3D plastic, rubber, foams, textiles), since the fractions differ in combustion and flight behavior and therefore influence cement quality. For this, training, validation, and test data were obtained from 11,526 manually sorted RDF particles, sampled from various German cement plants and processed using an at-line conveyor belt setup. The NIRS data were processed using a small convolutional neural network (CNN) to provide the respective fraction, yielding an accuracy of 99.5 %. The images were processed with different CNNs with transfer learning, yielding an accuracy of 96.7 %. In a second phase, both NIRS and image predictions were combined by soft voting, yielding an accuracy of 99.7 %. This validates the method under lab conditions and lays the groundwork for an application in a cement plant.

Keywords: Classification · Refuse derived fuel · Cement production · Near-infrared · CNN · Transfer learning

1 Introduction

1.1 Cement production and Refuse derived fuels

Cement is the most used building material worldwide [5]. Its main component, cement clinker, is produced from limestone, $CaCO_3$. Clinker consists of a mixture of CaO minerals, which give cement its characteristic properties. From the

chemical formulas of clinker formation it is obvious that producing clinker requires the removal of carbon dioxide from the limestone. This happens in the calcination reaction, shown in Eq. (1).

$$CaCO_3(s) \rightarrow CaO(s) + CO_2(g), \Delta h_r = +1780 kJ/kg \quad (1)$$

The reaction enthalpy of +1780 kJ/kg indicates that the reaction is highly endothermic. Consequently, the calcination reaction occurs at 900 °C [2]. Combined with other necessary reactions and heat losses, the mean specific thermal energy demand in the German cement production is 2807 MJ per 1 ton of cement [28]. The required heat is provided by direct combustion in a horizontal rotary kiln (diameter 3-6 m, length up to 100 m). The raw limestone (ground into a fine meal) is conveyed by gravity due to a kiln incline of approx. 5 % and a kiln rotation of ca. 1-4 rpm [14]. The fuels are injected together with the combustion air into the kiln at the main burner in counter flow to the raw limestone. The combustion of the fuels provides temperatures of 1450 °C near the flame [14].

The main emission of the cement production is CO_2, both from the fuels and the calcination reaction. An average German cement plant emits 586 kg of CO_2 per ton of cement [27]. While the emissions of the calcination are inevitable, the emissions of the fuel are directly influenced by the type of fuel used. Traditionally, fossil fuels like lignite and pulverized coal are combusted. This comes with several disadvantages: First, the fuels are expensive, finite and not locally available. Second, fossil fuels increase the carbon dioxide emissions of the cement production, which also, in countries with CO_2 certificates or tax further increase the fuel price. Third, the higher CO_2 emissions result in the acceleration of anthropogenic climate change. It is estimated that cement production is responsible for around 8 % of worldwide CO_2 emissions [13]. 60 % of these emission originate from the calcination reaction itself, but the remaining 40 % originate from the energy supply, mainly from the fuels [28].

Alternative fuels can help to mitigate these disadvantages. These fuels can be used tires, sewage sludge or animal bone meal, but the biggest proportion are refuse derived fuels (RDF), made from industrial and municipal waste. RDF are produced from waste in several steps, mainly involving shredding to sizes of few centimeters and removal of metals. RDF are not only cheap, in some cases they have a negative price (considered as waste incineration), which helps stabilizing production prices. In terms of CO_2 emissions certificates, RDF are also cheaper, since they 1) have a better carbon to hydrogen ratio and 2) contain a certain amount of biodegradable waste, which is considered CO_2 neutral. In total, RDF produces only around 50 g CO_2 per provided MJ energy [18] (lignite: 111 gCO_2/MJ [11], coal: 94 gCO_2/MJ [11]). Since industrial and municipal waste are generated at all urban and industrial areas, RDF do not need long and expensive routes of transport. The substitution ratios of alternative fuels in cement production differ worldwide: Austria (79 % [22]) and Germany (73 % [28]) are in pole position, the EU average is 46 %, while the United States only use 15 % and Asian countries like China, South Korea and Japan use 11 % [22]. In India, the substitution ratio is minimal, with a percentage of only 3%

[22]. Since the waste input streams differ over time and location, so does the RDF composition. The composition can be divided in six fractions: Paper and cardboard, foils, three-dimensional plastics, foams, textiles, and rubber. These fractions differ both in flight and combustion behavior and are therefore not equally suitable for the combustion in a cement plant. 3D plastics, for example, often hit the clinker bed not completely combusted, which produces reducing conditions in the clinker bed and therefore lower cement quality [1].

1.2 Near-infrared spectroscopy

Near-infrared spectroscopy (NIRS) is a method to determine the structure of molecules. It is a type of vibrational spectroscopy in which molecules are exposed to electromagnetic radiation. Depending on their structure, different vibrational modes (e.g., stretching, bending, or twisting) are possible. For energy in the near-infrared (NIR) wavelength range to be absorbed, a molecule must exhibit a dipole moment [16,23]. This makes NIRS particularly useful for detecting molecular bonds such as C=O, O-H, and C-H.

A typical NIRS setup consists of two main components. First, a light source emitting radiation in the NIRd wavelength range of 800–2500 nm, such as halogen lamps, LEDs, or lasers. The light interacts with a sample and is reflected from the sample. Second, the reflected light is detected by an NIRS detector, such as a silicon photodiode, photomultiplier tube, or an indium gallium arsenide (InGaAs) sensor, which differ in their detection wavelength ranges. The change in light intensity due to light absorption in the sample (measured as dimensionless absorbance) is then plotted as a function of wavelength, resulting in a characteristic spectrum [16,23].

Unlike mid-infrared (MIR) spectroscopy, where fundamental molecular vibrations are dominant, NIRS spectra primarily consist of overtones and combination bands. Overtones are harmonic vibrations that occur at integer multiples of the fundamental vibrational frequency [16]. In NIRS, the most prominent overtones originate from C-H, O-H, and N-H bonds, making the technique particularly sensitive to organic compounds and water.

For molecules with more than two atoms ($N > 2$), the number of vibrational degrees of freedom is given by $3N - 6$ [16]. As a result, NIR spectra can be complex, with overlapping absorption bands. While characteristic peaks can be associated with specific molecular bonds, manual classification based solely on NIR spectra is challenging. Therefore, statistical and machine learning methods are essential for the automated classification and analysis of NIRS data.

1.3 Main Contributions

The literature on classification and evaluation of RDF in the cement process is limited. The approach of fuel classification through proximate analysis and hierarchical learning [7] still relies on slow and labor-intensive laboratory analyses. Several methods for RDF classification using NIRS exist, but they lack proper validation: In [29], a classifier applied to pure substances (e.g., PET) was used

to predict RDF heating values. Since the fuel was not manually sorted, true labels were unavailable, meaning the accuracy and effectiveness of the classifier in predicting RDF fractions could not be verified. In Sevcik [21], FT-NIRS was combined with a Support Vector Machine, but only a small number of artificial model fuel samples were used (2-8 per fraction). Krämer and Flamme [12] report a commercial product for real-time RDF heating value analysis, but neither its accuracy nor the underlying algorithms are specified. For RDF classification using RGB images, an initial approach is proposed in Peddireddy et al. [17], but it still depends on manual sorting, essentially shifting the sorting process from a conveyor belt to a desk-based workstation. Additionally, a study from Tahir et al. [24] presents a method to detect and classify waste based on video recordings taken after processing municipal solid waste (MSW) through a mechanical sorting line. The method demonstrates an accuracy of 0.70. In Fischer et al. [9], a similar approach for RDF classification based on images showed an accuracy of 0.71. A combination of image and NIRS classification for RDF classification is not yet known.

2 Experiments

2.1 Experimental Setup

The experimental setup is designed as either an offline or at-line method to classify individual RDF particles and hence determine the composition of a fuel. The setup is constructed around a conveyor belt. An image is shown in Fig. 1. A camera, Basler a2A1920-160ucPRO with 2.3 MP and 168 fps, is mounted on top of the conveyor belt, such that the recorded images are not distorted by an angle. Laterally, both an infrared light source, a tungsten halogen lamp, and an near-infrared detector, Viavi 1700 ES, are mounted at an angle to the conveyor belt. The NIRS detector is an indium gallium arsenide detector with measurable wavelength in the range of 908 - 1676 nm. The measurement time is 9.2 ms and the wavelength resolution is 6.19 nm. To provide rich meta data, an environment sensor 2JCIE-BU01 from OMRON is used. It can be connected by USB 3.0 and measures temperature, humidity, illuminance, barometric pressure, sound noise, 3-axis acceleration, and TVOC (Total volatile organic compounds). Accuracies are given in [15].

The particles are fed by vibration over a 8 mm sieve to the conveyor belt. This ensures both no fines and dust on the conveyor belt, and separation of individual particles, such that they do not overlap and can clearly be distinguished from another. A measurement is triggered by a color change of central pixels against an average background of the black conveyor belt. For each measurement, one image, one NIR spectrum and one set of meta-data are stored into a SQL database. Each particle is assigned an unique ID (UUID4) and each experiment is grouped into a batch of particles. For the database creation, RDF mixtures from different cement plants were beforehand manually sorted into batches of the six fractions. The manual sorting was executed due to a pre-determined sorting catalog which defines optic, material and haptic of the particles. The presorted RDF particles

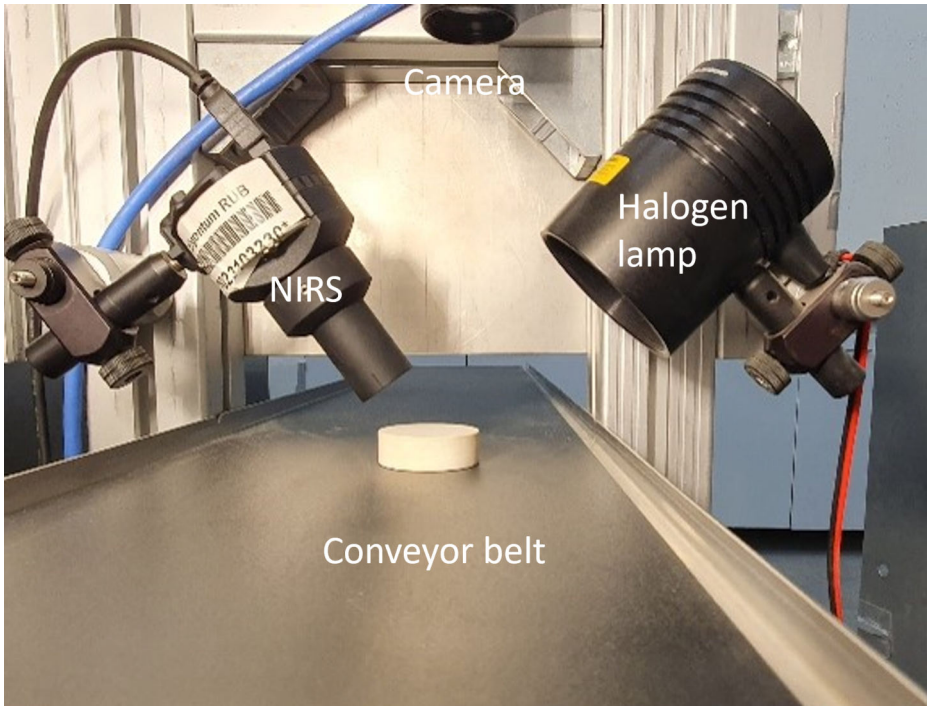


Fig. 1: Experimental setup consisting of camera (middle up) over a conveyor belt (middle) and a NIRS (left) with a halogen lamp (right).

were then fed to the experimental setup, hence, the batch labels could later be used as labels for classification.

2.2 Experimental dataset

In Table 1, the numbers of particles per fraction in the database are listed, resulting in a total number of 11,526. Since the fractions in RDF mixtures do not occur equally, it follows that the numbers per fraction in the database are not evenly distributed either. Rubber for example is rarer than other fractions, resulting only in 1066 samples, while textiles with 2580 samples are slightly oversampled. In Fig. 2, example images for each fraction are shown. Here, some problematic characteristics of RDF for quality control can be seen: RDF particles differ in size, tend to stick together and sometimes are black or transparent. Also, the particles are not always perfectly centered in the image. In Fig. 3, example NIR spectra for each fraction are depicted, showing typical behavior: While both plastic fractions, 3D plastics and foils, have distinct peaks around 1200 nm and 1450 nm, the fractions paper and cardboard (PC) and textiles show one broad peak from 1450 nm to 1650 nm. Rubber and foams show only smaller peaks. Especially for rubber, this is the case due to the dominant black color of the particles, which leads to an higher NIRS absorbance regardless of the material.

The dataset is published at `https://doi.org/10.5281/zenodo.14859683` [8]. For all following classification tasks, 10 % of each fraction were used as each validation and test data, leaving 80 % for the training data. In a first step, image and NIRS classification were executed separately, while in a second step, the results were combined. The implementation of neural networks was realized with TensorFlow [26]. The training was executed on a NVIDIA GeForce RTX 4070 Ti Super GPU (VRAM = 16 GB).

Table 1: Number of samples per fraction.

Fraction	**Number**
Foils	2188
3D plastics	1321
Paper and Cardboard	2442
Rubber	1066
Foams	1929
Textiles	2580
Total	11526

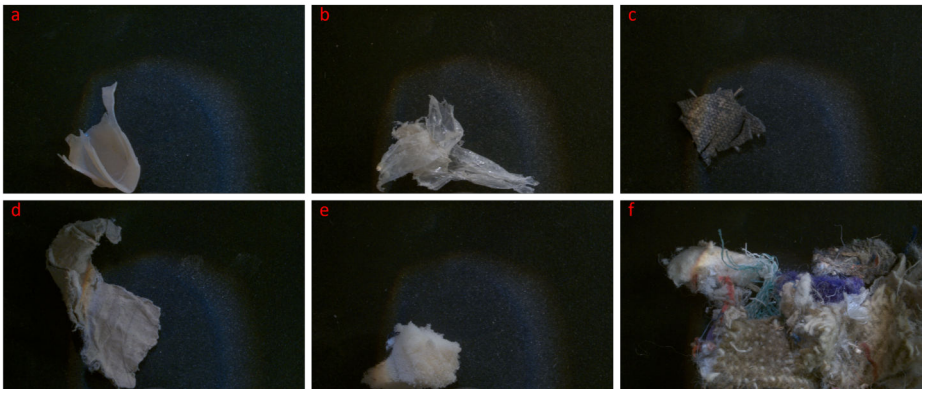


Fig. 2: Examples of RDF images per fraction: a 3D plastic, b foils, c rubber, d paper and cardboard, e foams, f textiles.

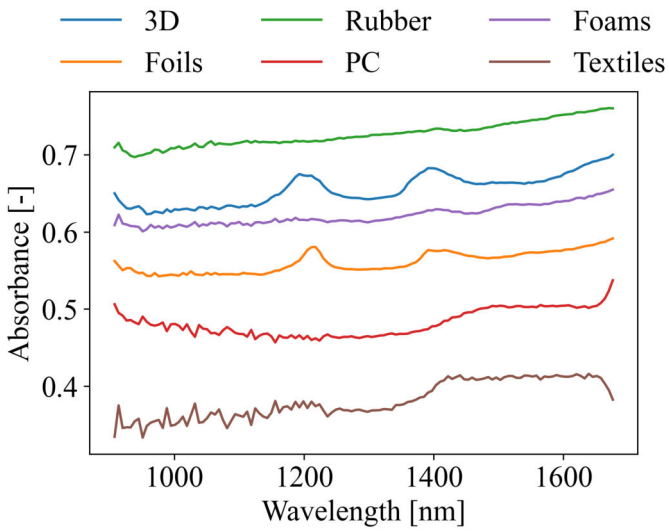


Fig. 3: Examples of NIR spectra per fraction.

2.3 Image classification

The images of the training data set were used together with the labels, inherited from the batch identities, to classify the samples in one of the six different fractions. For this, different convolutional neural networks (CNNs) were compared, namely Xception [4], ResNet-50 [10] and EfficientNets B0-B5 [25]. An overview is given in Table 2.

Additionally to this CNNs, from here on called "base models", some other layers were added, such that the data flow is as follows:

1. Input layer with size according to base model between 224x224x3 and 456x456x3 pixel.
2. Data augmentation layer with random flip, random translation, random rotation, random zoom, random contrast and random brightness.
3. Normalization layer.
4. Base model CNN.
5. Global average pooling layer.
6. Dropout Layer (0.5).
7. Batch normalization layer.
8. Dense connected layer with 128 nodes.
9. Output layer with softmax activation.

To increase training speed, all CNNs were trained with transfer learning, using the weights from the ImageNet dataset [19]. In transfer learning, the training is split into two training phases. The first phase is used to only train the dense connected layer at the end of the network with a high learning rate ($\alpha = 1e^{-3}$). In the second step, the whole network is trained, but with a smaller learning rate ($\alpha = 1e^{-5}$). In this publication, both phases lasted 50 epochs. The optimizer used is Adam with categorical crossentropy as loss function. The batch size was selected as 16, except for EfficientNetB5 as base model, where the batch size needed to be reduced to 8 due to a lack of more VRAM. Bigger EfficientNets therefore were not tested.

2.4 NIRS classification

For the classification based on NIRS measurement, the feature is the absorbance over the wavelength after some preprocessing (125 data points per sample), which is described in the following. After being loaded from the data base, a Savitzky-Golay filter [20] is applied to each spectrum. This is a common practice in chemometrics and it serves two purposes: First, smoothing of the data, reducing the measurement noise and second, a robust way to find a smooth derivative. Derivatives can help increase classification results since mostly peaks of absorbance are used to determine the identity of a sample. The filter is a finite filter with kernel length K_{len}, applied to the spectrum values x at position a by multiplying with kernel values K_r, as shown in Eq. (2).

Table 2: Used CNN based models with trainable parameters (including upstream and downstream), input size, ImageNet Top1 and Top3 accuracies.

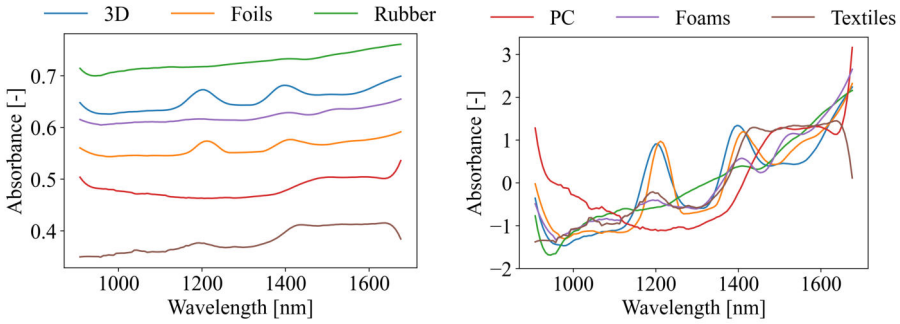
Base model	Trainable Parameters	Input size	ImageNet Acc. Top1	ImageNet Acc. Top3	Source
B0	4,219,433	224x224x3	0.771	0.933	[25]
B1	6,745,101	240x240x3	0.791	0.944	[25]
B2	7,955,327	260x260x3	0.801	0.949	[25]
B3	10,987,189	300x300x3	0.816	0.957	[25]
B4	17,911,269	380x380x3	0.829	0.964	[25]
B5	28,784,765	456x456x3	0.836	0.967	[25]
ResNet50	23,858,950	224x224x3	0.7715	0.933	[10]
Xception	21,132,718	229x229x3	0.79	0.945	[4]

$$\bar{x_a} = \sum_{r=1}^{K_{len}} K_r x_{r-a} \tag{2}$$

K_r depend on both K_{len} and the derivative d. Also, $d = 0$ is possible, which only applies a smoothing function to the measurement values. In Fig. 4 (a), the example spectra from Fig. 3 are shown after the application of the Savitzky-Golay filter ($K_{len} = 15$, $d = 0$). Since spectral data are heavily dependent on the lighting situation, the default mean value is not equal, such that a normalization is necessary to make NIRS measurements comparable and useful for classification. In this contribution, normalization is done with the standard normal variate (SNV). With SNV, the spectrum is normalized as shown in Eq. (3). First, the mean value of the spectrum $\bar{x}$ is subtracted. This leads to a new mean of 0. Second, the difference is divided by the standard deviation s. In this way, the variance σ is equal 1.

$$x_{SNV} = \frac{x - \bar{x}}{s} \tag{3}$$

The result is the feature used for the NIRS classification with mean of 0 and variance equal 1. In Fig. 4 (b), the spectra from Fig. 3 with SNV normalization are shown. The labels are, analogous to the image classification, derived from the batch identities. In Fig. 5, the structure of the used neural network for NIRS classification is shown. In the upper half, two blocks of 1D convolutional layers combined with batch normalization and maximum pooling process the spectrum, similar to [3]. An 1D convolution is, like the Savitzky-Golay filter, a finite filter, but with learned kernel values K_r. In this neural network, the 1D convolutional layers apply 64 filters each with a kernel size of 5. The second half consists of three dense connected layers with 128, 32 and 6 nodes. The last node activates with the softmax function, while all other layers activate with ReLu. In between of the blocks, Gaussian noise is added as data augmentation to increase robustness against measurement noise. Additionally, a dropout of 0.5



(a) (b)

Fig. 4: Data preprocessing of NIRS data example from Fig. 3. Left: Savitzky-Golay filter with $K_{len} = 15$ and $d = 0$ applied to example spectra. Right: Standard normal variate applied to results from left.

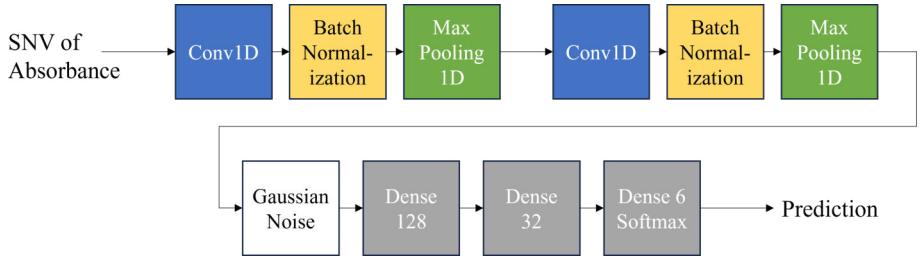


Fig. 5: Neural network structure for NIRS classification. The first half consists of two time, 1D convolutional layer, batch normalization and and maximum pooling. The second half consist of dense connected layers. During training, Gaussian noise added in between.

is used during training to prevent overfitting. The structure adds up to 255270 trainable parameters.

2.5 Combined image and NIRS classification

Combining predictions from different classifiers can benefit the accuracy of the prediction. For combining the predictions c_i^j of sample i and classifier j, different options are possible. In hard voting, the class with most predictions votes c_i^j is voted. When using only two classifiers, this would either result in an unanimous vote or a tie, which cannot increase the accuracy. Another form of voting is soft voting [31], where the prediction probabilities p_i^j are used as shown in Eq. (4).

$$p_i = \sum_{j=1}^{C} w_j p_i^j \tag{4}$$

C is the number of classifiers and w_j are weights. In this contribution, both the unweighted case ($w_{j=1,2} = 1$) as well as the weighted case are tested. For the weighted case, the weight of the classifiers are their accuracy on the test data. In majority voting, the prediction with the highest probability is used, as shown in Eq. (5).

$$c_i = argmax(p_i^1, ..., p_i^j) \tag{5}$$

3 Results and Discussion

3.1 Image classification

In Fig. 6 (a), the accuracies of the image classification with different base models are shown over the number of trainable parameters of the whole network with all upstream and downstream layers. First, it can easily be seen that both Xception and ResNet, although having a high number of trainable parameters, do not perform better than some smaller base models tested (EfficientNet B0-B3). For the EfficientNet base models, a linear correlation between trainable parameters and accuracy can be stated. In Fig. 6 (b), the accuracies are plotted over the corresponding training time. Surprisingly here, ResNet is the base model with the lowest training time, but differences to Xception an EfficientNets B0-B3 are small. For the EfficientNets models, there is again a quasi linear correlation between accuracy and training time as seen between training parameters and training time. Highest accuracies can be observed with base models EfficientNet B4 (accuracy=0.967) and B5 (accuracy=0.979). Although B5 provides the highest accuracy, training time is nearly double (7.1 h instead of 4.8 h) of B4, and can only barely be executed on the employed hardware. Hence, EfficientNet B4 is selected as base model for the image classification part. In Table 3, the confusion matrix for the image classification is listed. Foils and 3D plastics are often confused due to their optical similarity. Paper/cardboard has the lowest precision of 93.85 %, which results in mispredicted samples for all fractions except rubber. This fraction has a recall of 100 %, meaning only rubber samples were predicted as rubber. To measure the effect of transfer learning, the training with EfficientNet B4 was again executed with random initial weights. The training plan remained the same (50 epochs $\alpha = 1e^{-3}$, 50 epochs $\alpha = 1e^{-5}$), except that in the first phase also the whole model was trained. The training took 7.3 h, which is 1.5 times the required training time for the EfficientNet B4 using transfer learning. Although the accuracy on training and test data was high after the training, the accuracy on the test data was just 0.27, only slightly better than random. This shows the benefits of the transfer learning technique.

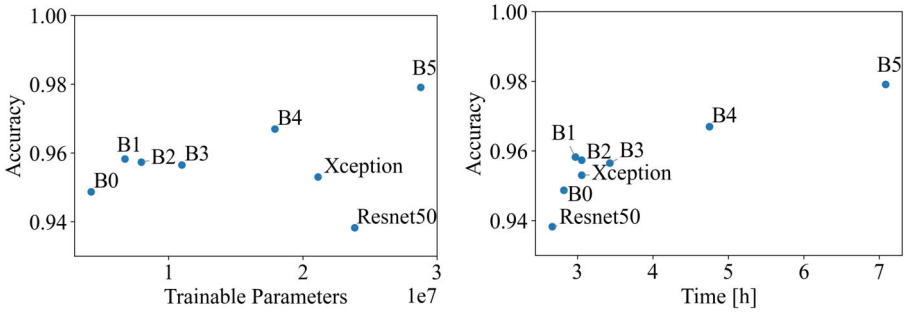


Fig. 6: Accuracies on the test data set of Xception, ResNet 50 and Efficientnet B0-B5 over (a) the number of trainable parameters (left) and (b) over the required training time in hours.

Table 3: Confusion Matrix for image classification.

		Predictions					
		3D Plastics	Foils	Rubber	Paper	Foams	Textiles
Identity	3D Plastics	0.9621	0.0303	0.0	0.0076	0.0	0.0
	Foils	0.0138	0.9725	0.0	0.0092	0.0	0.0046
	Rubber	0.0094	0.0	0.9623	0.0	0.0189	0.0094
	Paper	0.0164	0.0164	0.0	0.9385	0.0205	0.0082
	Foams	0.0	0.0052	0.0	0.0052	0.9896	0.0
	Textiles	0.0	0.0	0.0	0.0	0.0233	0.9767

3.2 NIRS classification

The accuracies for NIRS classification with derivatives $d = 0, 1, 2$ are 0.995, 0.724 and 0.763, respectively. Surprisingly, the results for $d = 0$ with an accuracy of 0.995 are the highest. This might be due to the convolutional layers in the used neural network, which, in combination with the smoothing Savitzky-Golay filter and SNV normalization, extract the required information to distinguish the different fractions. When using $d = 0$, not only peaks in the spectra are used for the prediction, but the whole shape and absolute values of the curve. Both differ drastically between two groups: While the 3D plastics and foil fractions show narrow, sharp peaks, the other fractions show broad peaks over many wavelengths (see Fig. 3). In Table 4, the confusion matrix for the NIRS classification is shown. For the fractions foils, 3D plastic and textiles the precisions are 100 % and for textiles also the recall is 100 %. Between 3D plastic and foils no mix-up occurred.

In Fig. 7, the feature importance by permutation is shown. For this, the feature was permuted 100 times for each wavelength with a kernel size of 5 (equal

Table 4: Confusion Matrix for NIRS classification.

		Predictions					
		3D Plastics	Foils	Rubber	Paper	Foams	Textiles
Identity	3D Plastics	1.00	0.0	0.0	0.0	0.0	0.0
	Foils	0.0	1.00	0.0	0.0	0.0	0.0
	Rubber	0.0189	0.0189	0.9623	0.0	0.0	0.0
	Paper	0.0041	0.0	0.0	0.9959	0.0	0.0
	Foams	0.0	0.0	0.0	0.0052	0.9948	0.0
	Textiles	0.0	0.0	0.0	0.0	0.0	1.00

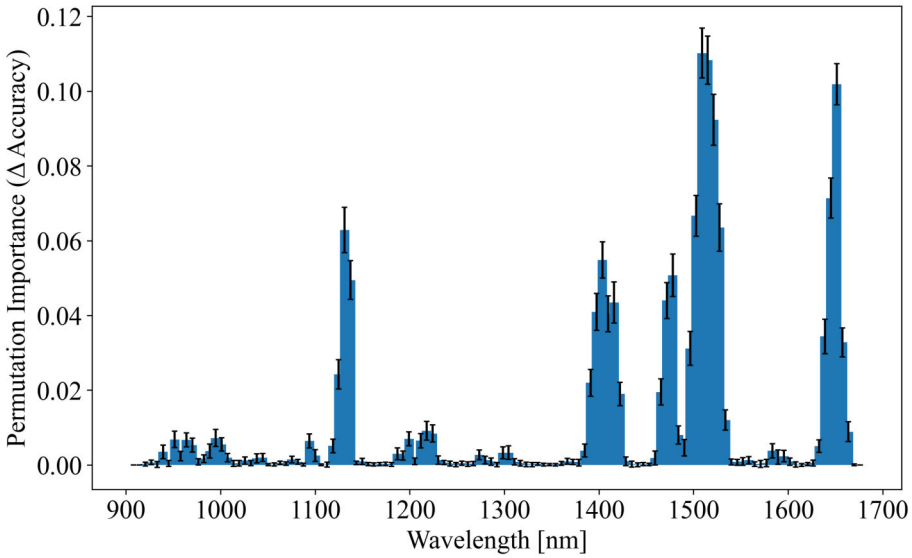


Fig. 7: Feature importance by permutation.

to the 1D Conv. layer kernel). This means, for each wavelength, the two adjacent values above and below were also permuted. The highest feature importance is observed at wavelength 1650 nm, which represents the rising absorbance at the upper end of the spectra. This is likely due to an olefinic C-H overtone at 1680 nm or an aromatic C-H overtone at 1685 nm [23]. Another region of high feature importance is the region around 1500 nm, which is probably due to the broad peaks of the textile and paper / cardboard fractions corresponding to O-H overtones [23] from cellulose [6]. The peaks at 1400 nm and 1200 nm correspond to the peaks of foils and 3D plastics (see Fig. 3), and their C-H (1200 nm) and O-H groups (1400 nm) [23]. The peak at 1130 nm may be the third overtone of C-H groups [30], for example seen in a small absorbance peak of textile particles.

Table 5: Confusion Matrix for combined classification.

		Predictions					
		3D Plastics	Foils	Rubber	Paper	Foams	Textiles
Identity	3D Plastics	1.0000	0.0	0.0	0.0	0.0	0.0
	Foils	0.0	1.0000	0.0	0.0	0.0	0.0
	Rubber	0.0094	0.0189	0.9717	0.0	0.0	0.0
	Paper	0.0	0.0	0.0	1.0000	0.0	0.0
	Foams	0.0	0.0	0.0	0.0	1.0000	0.0
	Textiles	0.0	0.0	0.0	0.0	0.0	1.0000

3.3 Combined image and NIRS classification

The accuracies for unweighted soft voting, weighted soft voting and majority voting are listed were 0.997, 0.996 and 0.997. The differences between these methods are only small, but the unweighted soft voting combination and majority voting provided both the best results with an accuracy of 0.997.

In Table 5, the confusion matrix for the combined classification with unweighted soft voting is shown. Only the predictions for some rubber samples are incorrectly predicted as 3D plastic or foils, but all other predictions are correct. Majority voting showed the same results with the same samples being predicted incorrectly. The remaining incorrectly predicted samples are three rubber particles, one black particle and two off-centered particles, shown in Fig. 8. All three samples were also falsely predicted with NIRS, while being predicted correctly with images. All of the falsely predicted samples of the image classification were predicted correctly with NIRS and vice versa. Each sample was therefore at least once predicted correctly.

Another approach, which can increase the robustness of the method is the application of a threshold t: Only predictions with a $p_i^j > t$ are admitted for the final prediction, otherwise the prediction is either only based on the other method or the sample has to be rejected in order not to compromise the statement about the composition. For $t = 0.9$, 96.2 % of all predictions from the EfficientNet B4 and 99.9 % of all predictions from the NIRS CNN were above the threshold.

In comparison to reported accuracies in literature (around 0.7, see section 1.3), the accuracies for both image and NIRS classification separate and combined are higher. In comparison to these approaches, this contribution uses a bigger and manual labeled data set.

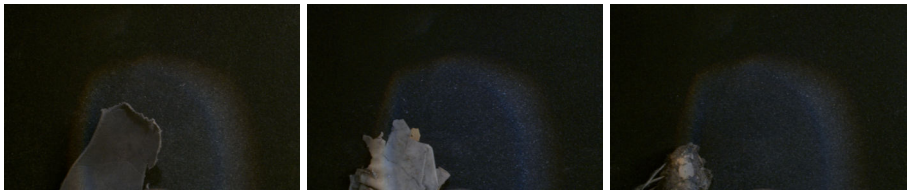


Fig. 8: Samples incorrectly predicted with unweighted soft voting.

4 Conclusion

Two different measurement techniques were used to characterize refuse derived fuels (RDF), RGB images and near-infrared spectroscopy (NIRS). A dataset with 11526 samples was created with samples from different cement plants on an lab set-up. The dataset was then used to classify RDF in one of six fractions: paper, foils, 3D plastic, rubber, foams, textiles. In a first step, image and NIRS classification were considered separately. For image classification, transfer learning for different CNNs with pre-trained ImageNet weights was tested. The highest accuracy in proportions to training effort and trainable parameters was achieved with EfficientNets. For use in the combined predictor, EfficientNet B4, which showed an accuracy of 0.967 on the test data, was selected. For the NIRS classification, a self-designed CNN with 1D convolutional layers was used. Surprisingly, the CNN showed better results when using the spectrum directly instead of a derivative of the spectrum. This predictor showed an accuracy of 0.995 on test data. To combine both methods in a second step, soft voting and majority voting were tested, which increased the accuracy to 0.997. With this, the combined technique is lab validated. Nevertheless, a validation in a cement plant is the next required step. For example, influences of the harsh environment in a cement plant (vibration, dust, humidity) need to be examined.

Acknowledgments. The IGF Project (01IF22676N) is supported within the programme for promoting the Industrial Collective Research (IGF) of the German Ministry of Economic Affairs and Energy.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bodendiek, N., van Thriel, H., Schäfer, S., Hoenig, V., Wirtz, S., Scherer, V.: Untersuchung der Wechselwirkung von Brennstoffpartikeln und Zementklinker zur Erhöhung des Ersatzbrennstoff-Einsatzes: 30. Deutscher Flammentag (2021)
2. Chatterjee, A.K.: Cement Production Technology: Principles and Practice. Chapman and Hall/CRC, Milton (2018), https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=5372003

3. Chen, X., Cheng, G., Liu, S., Meng, S., Jiao, Y., Zhang, W., Liang, J., Zhang, W., Wang, B., Xu, X., Xu, J.: Probing 1D convolutional neural network adapted to near-infrared spectroscopy for efficient classification of mixed fish. Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy **279**, 121350 (2022). https://doi.org/10.1016/j.saa.2022.121350
4. Chollet, F.: Xception: Deep Learning With Depthwise Separable Convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2017)
5. Crow, J.M.: The concrete conundrum. Chemistry World (2008), https://www.rsc.org/images/Construction_tcm18-114530.pdf
6. dos Santos, E.O., Silva, A.M.S., Fragoso, W.D., Pasquini, C., Pimentel, M.F.: Determination of degree of polymerization of insulating paper using near infrared spectroscopy and multivariate calibration. Vibrational Spectroscopy **52**(2), 154–157 (2010). https://doi.org/10.1016/j.vibspec.2009.12.004
7. Elmaz, F., Büyükçakır, B., Yücel, Ö., Mutlu, A.Y.: Classification of solid fuels with machine learning. Fuel **266**, 117066 (2020). https://doi.org/10.1016/j.fuel.2020.117066
8. Fischer, J., Fehler, L., Treiber, K., Scherer, V.: Near-Infrared Spectroscopy and Image Classification of Refuse-Derived Fuels for Cement Production. https://doi.org/10.5281/zenodo.14859682
9. Fischer, J., Wirtz, S., Scherer, V.: Random forest classifier and neural network for fraction identification of refuse-derived fuel images. Fuel **341**, 127712 (2023). https://doi.org/10.1016/j.fuel.2023.127712
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition, http://arxiv.org/pdf/1512.03385
11. Juhrich, K.: CO2-Emissionsfaktoren für fossile Brennstoffe: CLIMATE CHANGE 27/2016
12. Krämer, P., Flamme, S.: Real-time Analysis of Solid Recovered Fuels using Sensor Technology. In: Thomé-Kozmiensky, K.J., Thiel, S. (eds.) Waste Management, pp. 339–348. Thomé-Kozmiensky Verlag GmbH, Nietwerder (2015)
13. Lehne, J., Preston, F.: Making Concrete Change: Innovation in Low-carbon Cement and Concrete, https://www.chathamhouse.org/sites/default/files/publications/research/2018-06-13-making-concrete-change-cement-lehne-preston.pdf
14. Locher, F.W.: Zement: Grundlagen der Herstellung und Verwendung. Verlag Bau+Technik, Düsseldorf (2000), https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=2029875
15. OMRON Corporation: Environment Sensor (USB Type) Environment Sensor (USB Type) 2JCIE-BU01: User's Manual, https://components.omron.com/eu-en/sites/components.omron.com.eu/files/ds_related_pdf/A279-E1.pdf#page=125.09
16. Ozaki, Y., Huck, C., Tsuchikawa, S., Engelsen, S.B. (eds.): Near-Infrared Spectroscopy. Springer Singapore, Singapore (2021). https://doi.org/10.1007/978-981-15-8648-4
17. Peddireddy, S., Longhurst, P.J., Wagland, S.T.: Characterising the composition of waste-derived fuels using a novel image analysis tool. Waste management (New York, N.Y.) **40**, 9–13 (2015). https://doi.org/10.1016/j.wasman.2015.03.015
18. Pohl, M., Becker, G., Heller, N., Birnstengel, B., Zotz, F.: Auswirkungen des nationalen Brennstoffemissionshandels auf die Abfallwirtschaft

19. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV) **115**(3), 211–252 (2015). https://doi.org/10.1007/s11263-015-0816-y
20. Savitzky, A., Golay, M.J.E.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry **36**(8), 1627–1639 (1964). https://doi.org/10.1021/ac60214a047
21. Ševčík, M.: Near-Infrared Spectroscopy for refuse derived fuel feasibility study of inorganic chlorine content quantification: Classification of waste material components using hyperspectral imaging, https://www.diva-portal.org/smash/record.jsf?pid=diva2:1281468
22. Sharma, P., Sheth, P.N., Mohapatra, B.N.: Recent Progress in Refuse Derived Fuel (RDF) Co-processing in Cement Production: Direct Firing in Kiln/Calciner vs Process Integration of RDF Gasification. Waste and Biomass Valorization **13**(11), 4347–4374 (2022). https://doi.org/10.1007/s12649-022-01840-8
23. Siesler, H.W. (ed.): Near-infrared spectroscopy: Principles, instruments, applications. Wiley-VCH, Weinheim, 2. reprint edn. (2005)
24. Tahir, J., Ahmad, R., Tian, Z.: Calorific value prediction models of processed refuse derived fuel 3 using ultimate analysis. Biofuels **14**(1), 69–78 (2023). https://doi.org/10.1080/17597269.2022.2116771
25. Tan, M., Le, Q.V.: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks (2019). https://doi.org/10.48550/arXiv.1905.11946
26. TensorFlow Developers: TensorFlow (2024). https://doi.org/10.5281/ZENODO.4724125
27. Verein Deutscher Zementwerke e.V.: Umweltdaten der deutschen Zementindustrie 2022, https://www.vdz-online.de/wissensportal/publikationen/umweltdaten-der-deutschen-zementindustrie-2022
28. Verein Deutscher Zementwerke e.V.: Umweltdaten der deutschen Zementindustrie 2023, https://www.vdz-online.de/wissensportal/publikationen/umweltdaten-der-deutschen-zementindustrie-2023
29. Verga, S., Compare, M., Zio, E., Carra, G., Farina, M., Righetto, I., Sala, V.: Estimation Of The Lower Heating Value Of Solid Recovered Fuel Based On Swir Hyper-Spectral Images And Machine Learning. In: 2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS). pp. 1–5. IEEE (2022). https://doi.org/10.1109/WHISPERS56178.2022.9955135
30. Weyer, L.G., Lo, S.C.: Spectra– Structure Correlations in the Near–Infrared. In: Griffiths, P.R., Chalmers, J.M. (eds.) Handbook of Vibrational Spectroscopy. Wiley (2001). https://doi.org/10.1002/0470027320.s4102
31. Zhou, Z.H. (ed.): Machine learning. Springer eBook Collection, Springer Singapore and Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-1967-3