

IMAGE-BASED DEEP LEARNING FOR RHEOLOGY DETERMINATION OF BINGHAM FLUIDS

Anne Ponick^{1*}, Amadeus Langer¹, Dries Beyer², Max Coenen², Michael Haist², Christian Heipke¹

¹ Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Germany
(ponick, langer, heipke)@ipi.uni-hannover.de

² Institute of Building Materials Science, Leibniz Universität Hannover, Germany
(d.beyer, m.coenen, haist)@baustoff.uni-hannover.de

Commission II, WG II/7

KEY WORDS: Rheology, Building Materials, Stereo View, Deep Learning

ABSTRACT:

In this work, a method to predict the rheological properties of ultrasonic gel, as a reference substance of cement paste, is presented. For this purpose, images are taken with a stereo camera system which show a mixing paddle moving through the ultrasonic gels of different consistency, thus setting them in motion. A digital elevation model (DEM) and a corresponding orthophoto are created from the image pairs using classical image matching and orthoprojection methods. These are used as inputs into a Convolutional Neural Network (CNN), which predicts the support points of a flow curve which classically have to be determined in a rheometer in the laboratory. A simple network architecture consisting of a small number of convolution layers is compared with a pre-trained ResNet-18, which is fine-tuned using gel images. In a second series of experiments, rheological parameters, which alternatively need to be deduced from the flow curve in a separate step, are determined directly from the images. In the third series of experiments, the influence of different factors is tested, such as the position of the cameras relative to the direction of paddle movement and the importance of the DEMs and orthophotos in the training. It is shown in this paper that it is possible to predict the rheological properties of the ultrasonic gels with a suitable setup with a satisfying accuracy.

1. INTRODUCTION

Concrete is one of the most widely used building materials in the construction industry; in Germany alone, the production volume of ready-mixed concrete in 2020 was over 50 million m³ (BTB, 2021). To conserve natural resources, demolition material (concrete, masonry, rock) from the deconstruction of buildings can be reused in concrete production in the form of recycled aggregates. In today's construction industry, however, mineral resources tend to be downcycled rather than recycled, although according to the German Federal Statistical Office, the construction sector was responsible for more than 55 % of the total German waste generated in 2019 (Statistisches Bundesamt (Destatis), 2021). One of the reasons is the fact that the properties of the demolition materials to be recycled vary widely and can negatively influence the new concrete to an unknown degree – here in particular the fresh concrete properties (workability, tendency to segregate). To overcome this problem, more cement can be used in the concrete mixing process when working with recycled aggregates to compensate for these fluctuations. However, this is a solution that is neither environmentally nor economically justifiable, because large amounts of CO₂ are released during the production of cement. If the properties of the freshly mixed concrete could already be determined during its mixing process, it would be possible to compensate the unknown negative effects, which mainly result from the use of recycled materials and from variations in the mixing process, e.g. with suitable additives.

The ReCyCONtrol¹ research project aims at achieving this goal. To this end, one part of the project is the observation of

the concrete mixing process using optical stereo sensors. The aim is to evaluate the flow behaviour of the concrete in real-time and, thus, to predict its rheological properties. Based on this data, the concrete can then be specifically adjusted towards the desired characteristics.

To describe the rheological properties of concrete, the Bingham model is used in most cases (Yahia et al., 2016), which is described by the plastic viscosity η in Pa·s and the yield stress τ_0 in Pa. These values describe the relationship between the shear rate $\dot{\gamma}$ in 1/s and the shear stress τ in Pa by

$$\tau = \tau_0 + \eta\dot{\gamma}. \quad (1)$$

The difference to a Newtonian fluid (such as water) is that a certain shear stress, the yield stress, must first be reached in order to set the substance into motion. Plastic viscosity, on the other hand, describes how flowable a substance is when it is in motion. While the plastic viscosity of fresh concrete can be roughly estimated visually by experts during the mixing process, this is not the case for the yield stress. In this work, Deep Learning is used to identify both quantities, as the network learns correlations between the values and the visual appearance itself without the need for expert knowledge.

As a first step and proof of concept, it is shown that it is possible to determine the rheological properties of Bingham fluids in motion from image data. Since working with fresh concrete always involves a lot of effort and since the time that can be worked with the fresh concrete before it hardens is limited, this work uses coloured ultrasonic gel mixed with water to varying degrees, a substitute which is a commonly used reference substance for cement paste (one of the main components of con-

* Corresponding author

¹ <https://www.recycontrol.uni-hannover.de/de/>

crete (Neroth and Vollenschaar, 2011)). Clay granulate is used to simulate the aggregate in the concrete. Thus, more tests can be carried out where different camera and illuminant positions are tested – especially with regard to 3D reconstruction.

In this paper, a learning procedure is presented that uses a digital elevation model (DEM) and an orthophoto, both derived by classical photogrammetric methods, to predict rheological properties. A regression is performed by a Convolutional Neural Network (CNN), which outputs specified values describing the rheological properties in each case. The main contribution of this paper is to prove the general possibility to extract rheological properties from images of Bingham fluids.

The paper next gives an overview of current research on automated processes in the construction industry. In section 3, the methodology used in this work is presented. The data set is described in section 4 including the data generation and the determination of the reference values. Finally, in section 5, the experiments are presented and the results are discussed. Section 6 concludes the paper and discusses future work.

2. RELATED WORKS

Automatic processes for the quality assurance of concrete are not widespread in the construction industry. Research in the automation of processes exists in the field of hard concrete, among others. (Song et al., 2020) present an approach to perform a semantic segmentation of concrete sections using Deep Learning to determine air voids. In (Coenen et al., 2021), a semi-supervised approach is used to segment the concrete aggregate, which can be used to determine the particle distribution. However, an important research point is whether or not the quality of a concrete can be predicted before it is used.

On the construction site, the quality of the fresh concrete is currently checked with classical manual methods, such as the slump test. By taking a stereo image of the "concrete cake", (Tuan et al., 2021) show that the manual slump measurements can be replaced by image measurements. The authors argue that human involvement requires time and effort and also has limited accuracy, while the proposed method, not requiring human involvement, has many potential advantages and ensures a better reliability of results. An automatic classification into slump cases via Deep Learning has also shown success. However, the concrete is only assessed in the mixed state and must be disposed of, if the quality is insufficient. Methods with which the quality can be determined before or during the mixing process are therefore of significant importance.

In (Chidiac and Mahmoodzadeh, 2009), the most common models of concrete technology and literature for determining the plastic viscosity of concrete based on its composition are reviewed. The results show that there are different ways predicting plastic viscosity from model to model and there are variations in the results, too. The most recent work predicting concrete properties based on its composition is (Nguyen et al., 2020). Using a hybridization of a Least Squares Support Vector Machine (LSSVM) and Particle Swarm Optimization (PSO), the yield stress is determined in addition to the plastic viscosity, leading to good performance. However, none of the models can handle the use of recycled aggregates, as their composition can vary considerably, leading to different water demands during the concrete mixing process for instance.

In (Lau Hiu Hoong et al., 2020), the problem of fluctuations in recycled aggregates is addressed and a near real-time method is developed to determine the composition of the recycled aggregate. The methodology is based on image analysis using a CNN. With the help of this method, the recycled aggregates can be pre-sorted based on their quality. It might be possible to extend the models for rheology prediction with the previously mentioned method, but uncertainties in the prediction are still to be expected due to fluctuations in the mixing process.

Even if the desired mixing ratios are known, inaccurate ratios can still occur in the mixing process. In (Yang et al., 2020), a method is developed to monitor this ratio by taking images of the fresh concrete immediately after mixing and using Deep Learning methods to perform a multilabel classification that assigns the concrete to one of five classes of water-to-binder-ratio and sand-to-aggregate-ratio, and one of three classes of nominal maximum particle size of coarse aggregate classes.

Both, (Li and An, 2014) and (Ding and An, 2018) evaluate images of concrete from the mixing process in a single-shaft mixer based on its workability. While (Li and An, 2014) use classical image analysis methods and determine the slump flow values and the V-funnel flow time by extracting the shape of the concrete in the mixer using pre-defined features, (Ding and An, 2018) show that it is also possible to determine the two values using Deep Learning, the method is thus independent of human experience and insights into the appearance of concrete with different rheological properties, once training data have been acquired.

The methods mentioned above usually only attempt to automate tests (e.g. the slump test) that are still carried out by human hands on the construction site. It has been shown that the related values can also be determined to a certain degree during the mixing process. Our aim is to predict the rheological properties from information acquired during the mixing process by automatically determining a flow curve from which values for the two rheological properties of yield stress and plastic viscosity are determined. Based on a reference flow curve from the laboratory the method is evaluated. In principle, it runs in real-time, thus, suitable additives can be added directly during the mixing process.

3. DEEP LEARNING FOR RHEOLOGICAL PROPERTY DETERMINATION

3.1 Introduction

Our goal is to develop a method that can predict the rheological properties of a Bingham fluid serving as reference substance for fresh concrete by observing the mixing process with imaging sensors. This method should then be adapted to real fresh concrete, however, this second step is beyond the scope of this paper. Using Deep Learning, the properties of the substance which simulates fresh concrete are to be determined via a regression model. A CNN learns features of recorded images, which show a mixing paddle moving through the reference substance at different time steps. Since it is assumed that additional information can be derived from the three-dimensional surface shape of the flowing substance, we use a stereo camera system in our approach.

It is of course possible to feed the network with the two stereo images as input. The network then has to implicitly learn the

3D surface information assumed to be relevant for the task to be solved. In a second scenario, the 3D information can be calculated beforehand and then used directly as input. Since in the first scenario less information is provided and the rest (here the DEM) consequently has to be learned, more training data tends to be needed. Given, that training data is scarce, the second scenario is investigated in this work. In a two-step procedure, first the 3D surface of the concrete is calculated from the captured stereo images using classical image matching methods. In order to co-register the images and the DEM, an orthophoto is then calculated from the input images. Consequently, the input of the CNN is a two channel image consisting of the orthophoto (as greyscale image) and the DEM. In our work photogrammetric processing is carried out using the commercial software *Agisoft Metashape*².

The outputs of the CNN are either the support points of the flow curve (9 output neurons in our case) or the regression parameters (2 output neurons) – calculated from the flow curve –, which have a direct link to two important quantities in concrete rheology, i.e. plastic viscosity and yield stress. Both approaches are tested because, although in most cases the main interest lies in the plastic viscosity and the yield stress, there are different methods to determine these from the flow curve, and thus an individual calculation can still be carried out when predicting the flow curve. Another point is that the flow curve contains additional information that can be analysed by building material experts, if desired.

In the following subsections, the used CNN architectures are described as well as the training procedure.

3.2 Network architectures

We use two different network architectures in this work: The first one is a simple architecture, here referred to as *Default CNN*, which consists of 7 convolution layer with a 5x5 kernel and a stride of 2, each followed by batch normalisation and ReLU as activation function. A fully connected layer maps the features to the 9 or the 2 output neurons. In total, this architecture has 377 457 and 376 106 parameters, respectively, which are estimated from scratch in the training process. This architecture was chosen because for a CNN it has relatively few parameters, which in itself is also a type of regularisation and can thus prevent overfitting to the training data.

Another method to deal with limited amounts of training data is to use a pre-trained network. The weights of these networks serve as good initial values and are adapted during re-training. For a first set of experiments, an existing CNN architecture pre-trained on the ImageNet (Deng et al., 2009) is used. However, the first layer and the last layer must be completely trained from a random initialisation, since a different number of input channels and of output neurons are needed. We are aware that the weights are pre-trained on RGB imagery, but the assumption is that the pre-trained weights can still be a support in the training process. Also, for this reason, all layers are re-trained, i.e. not only the last layers are fine-tuned.

In preliminary experiments, different architectures were tested, all with a small number of layers (ResNet-18 (He et al., 2016), VGG-11 (Simonyan and Zisserman, 2015), GoogLeNet (Szegedy et al., 2015) and AlexNet (Krizhevsky et al., 2012)). All of these architectures have shown very similar performance,

² <https://www.agisoft.com/>

which is why only the experiments with one architecture are presented in this paper; we have chosen ResNet-18. Having a 2-channel input and an output layer consisting of 9 neurons, the ResNet-18 architecture has a total of 11 177 993 trainable parameters, which is roughly 30 times more than what the *Default CNN* has.

3.3 Training

Training of the Network is carried out as follows. The network weights ω are iteratively adjusted by minimising an error function $E(\omega)$. In this work, the mean squared error (MSE) is used. For a batch of N samples, the squared difference of all K outputs \hat{y}_n^k of a sample x_n and the corresponding true values y^k are averaged over all outputs of all samples. This gives

$$E(\omega) = \frac{1}{N} \sum_{n=1}^N E_n(\omega, x_n) = \frac{1}{N \cdot K} \sum_{k=1}^K \sum_{n=1}^N (y_n^k - \hat{y}_n^k)^2. \quad (2)$$

As another tool to prevent over-fitting in addition to those mentioned so far, weight decay is used. It is added with a factor of λ to the error function and is a penalty term for large weights, leading to

$$E(\omega) \leftarrow E(\omega) + \lambda \sum_l \omega_l^2. \quad (3)$$

4. DATA GENERATION

4.1 Image acquisition

Ultrasonic gel is regularly used as a reference substance for cement paste (one of the main components of concrete) to investigate rheological properties (Haist et al., 2020). A major advantage of using ultrasonic gel is that its rheological properties – unlike cement paste – remain constant over a longer period of time, which means it can be used several times for experiments. In an experimental set-up, we currently use grey-coloured gel mixed with water to varying degrees. To simulate the aggregate in the concrete, clay granulate is added. The mixture is then filled into a horizontal channel and set in motion with a mixing paddle. This process is recorded in several runs with a stereo camera system. A graphical illustration of the set-up can be seen in Figure 1. The cameras used in this work are Grasshopper 3 USB cameras with a focal length of 8 mm. The 1920 px x 1200 px images were acquiesced at a frequency of 30 Hz. A trigger signal was sent from one camera to the other when a picture was taken. To check the synchronicity, an LED array was used in which 20 LEDs display a time in millisecond intervals. By combining the LEDs that were switched on and off, a time stamp was generated, which confirmed that the stereo image capture indeed occurred at the same instant in time. A challenge was the placement of the camera together with the illumination so that the ultrasonic gel would not reflect too strongly and thus, lead to errors in the 3D reconstruction and beyond. For this reason, the channel was completely darkened and two lights were placed in a way that they did not shine directly on the gel.

For each of the 21 mixed gel samples, we recorded at least 14 image sequences; the mixing paddle moved towards the cameras in 7 sequences, and in the other 7 sequences, the paddle

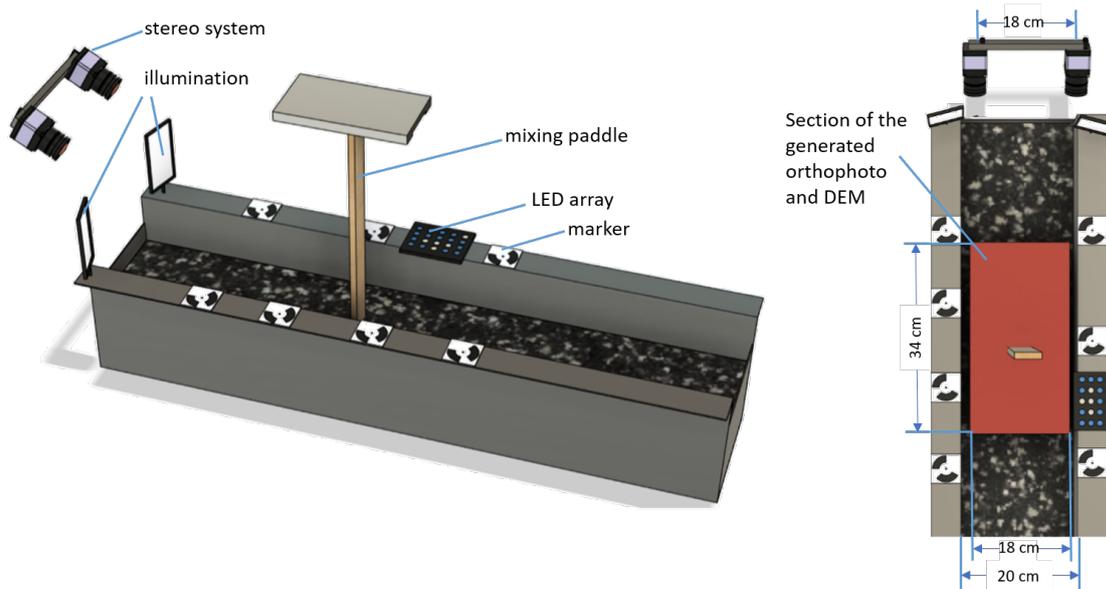


Figure 1. Experimental set-up for data generation. The geometric dimensions can be found in the top view (right).

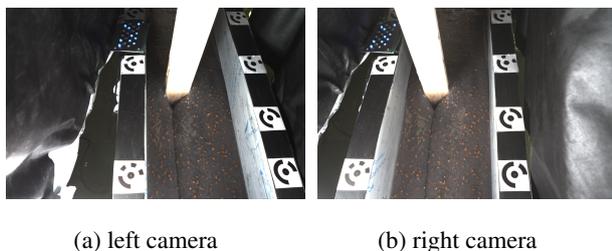


Figure 2. Pair of original images of sample 19, acquired at the same point of time. In the upper left part of each image, the blue LED panel can be seen.

moves away from the cameras. Each sequence consists of 40 image pairs. A DEM with an corresponding orthophoto of size 360 px x 679 px was then created from each stereo pair and was smoothed over 5 time steps over each image sequence with a box filter. By defining the datum of the 3D coordinates using stable markers located on the edge of the channel, the results of all pairs refer to an identical coordinate system. A challenge regarding the reconstruction were the occlusions caused by the mixing paddle. In Fig. 2, a pair of images is presented. Fig. 3 shows the depth maps and the orthophotos of three different samples of the ultrasonic gel. The images also depict the erroneous reconstruction in the orthophoto and the blurred depth in the depth image behind the mixing paddle caused by occlusion.

4.2 Reference values

As a reference, the different gel samples are placed in a viscometer in a laboratory, and a flow curve is determined through nine support points. The measurements were carried out with the *Viskomat NT* from *Schleibinger*³. 370 ml of each sample is placed in a rotatable round vessel and a paddle is immersed. The torque (Nmm) needed to obtain a certain rotational speed of the vessel (min^{-1}) is then measured at 9 different rotational speeds (50, 40, 30, 20, 10, 8, 6, 4, 2 min^{-1}) for 30 s each, starting at the highest speed. All samples should have been measured twice, but due to device failures in eight cases only one measurement was usable.

³ <http://www.schleibinger.com/>

Figure 4 shows the flow curves of the 21 samples and indicates the support points. It is noticeable that there is a clustering of flow curves with similar values in the middle of the value range and that the value gaps between the samples are larger towards the outside. From these measurements, the yield stress and the plastic viscosity can be derived.

For the experiments in which the yield stress and the plastic viscosity are to be determined directly, the related reference values are calculated from the flow curve using the method from (Haist et al., 2020). Here, a regression is calculated from the almost linear part of the flow curve (between 20 and 50 min^{-1}). With the multiplication of previously determined factors, the slope m in Nmm-min and the y-axis intersection n in Nmm of the regression line can be mapped to plastic viscosity in Pa-s and yield stress in Pa.

In table 1, the values of the support points and the corresponding regression parameters are listed. The difference to a second measurement, if it exists, is given in brackets. It can be seen that there were sometimes larger differences in a measurement for the same sample (e.g. sample 19 and sample 23). Especially for the eight samples of which there is only one measurement, it is difficult to assess, how much they can be trusted.

5. EXPERIMENTS

5.1 Rationale

In the experiments, the following aspects are investigated. First, a comparison between the *Default CNN* architecture with a random initialisation and the pre-trained *ResNet-18* is carried out for the prediction of the nine support points of the flow curve. In the second series of experiments, the regression parameters of the flow curve, which have a direct link to plastic viscosity and yield stress, are determined directly. These can then be compared with the regression parameters obtained by calculating a regression from the predictions of the support points from experiment 1. The third set of experiments deals with the influence of individual variables, either by working only with images in which the mixing paddle moves towards or only away

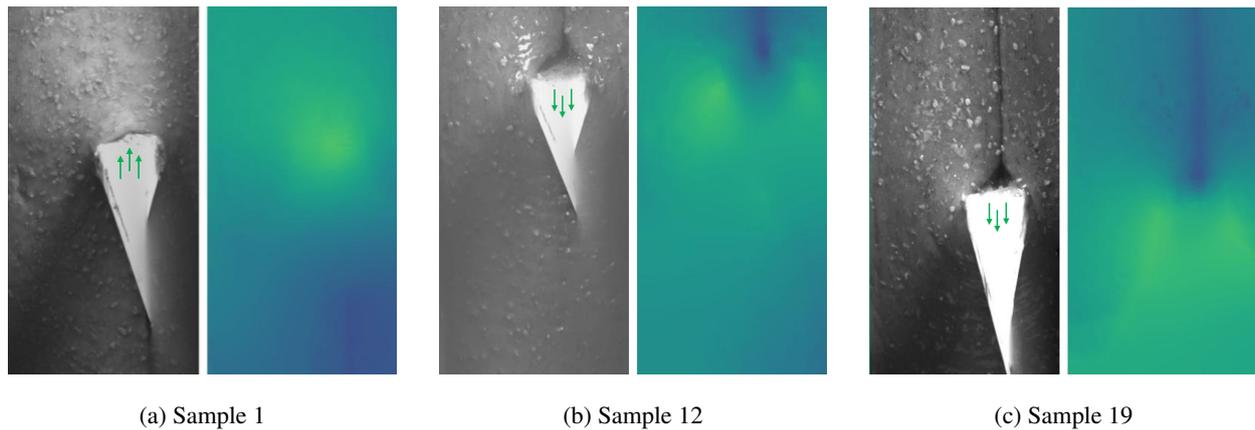


Figure 3. Examples of the input data: Generated orthophoto and depth map of different samples. The green arrows show the moving direction of the paddle. If the direction of movement is towards the camera (a), the DEM primarily shows a pile of gel that the mixing paddle creates in front of it. If the direction of movement of the paddle is away from the camera (b and c), the reconstruction mainly shows the groove that the paddle pulls through the gel. In the DEM, green corresponds to a higher and blue to a lower elevation.

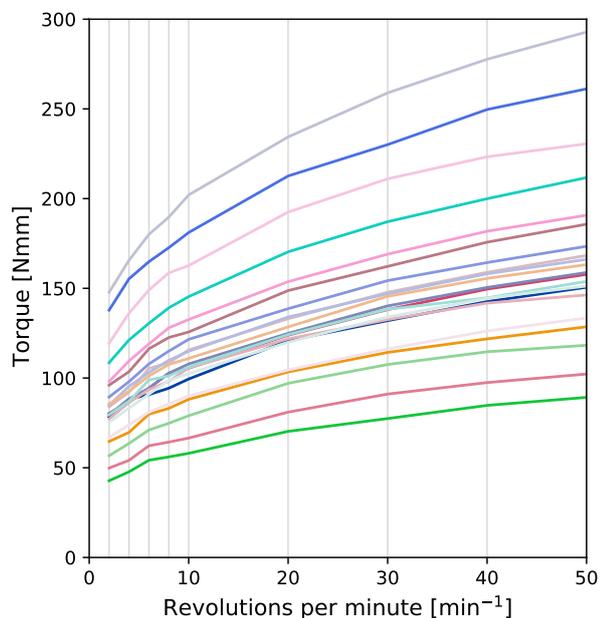


Figure 4. Plot of the flow curves generated in the rheometer. The different coloured lines show the measurements of the ultrasonic gel samples. The grey vertical lines show the points of the flow curve to be predicted.

from the camera. Especially for future use, it is interesting to see whether or not the network can generate more precise information about the behaviour of the substance in front of or behind the mixing paddle. In addition, the influence of the orthophoto and the DEM is tested by training with only one of the two inputs. The input for the CNN is considered to be individual independent images and/or DEMs without any connection in time. The fact that the input comes from image sequences is thus not exploited in the research reported here and will be investigated in future work.

5.2 Training configuration

The available samples are divided into training (15 samples), test and validation set (3 samples each). To do so, 3 splits are

randomly generated so that the evaluation can be done on different samples. In figure 6, the first column shows the distribution of the samples in the individual sets. The network training is performed with Stochastic Gradient Decent (SGD) using a Nesterov momentum of $\beta = 0.99$ based on the formula of (Sutskever et al., 2013). The learning rate for the *Default CNN* architecture is set to $1 \cdot 10^{-2}$, while a learning rate of $1 \cdot 10^{-3}$ is used for the *ResNet-18* architecture. A lower learning rate showed better results for the validation set in some preliminary tests with the pre-trained networks, probably because the weights are already rather accurate and do not need to be adjusted too much further. The weight decay parameter was set to $\lambda = 1 \cdot 10^{-3}$ for both architectures. These parameters were determined in preliminary tests. Training is carried out for a maximum of 1000 epochs (an epoch comprises one run of the complete training set), but is terminated prematurely if the validation accuracy has not improved for 250 epochs. The evaluation is then carried out with the parameters that showed the best accuracy during validation.

The grey values of the orthophotos are normalised per sample to mean 0 and standard deviation 1. In the DEMs, all images in a sequence are subtracted from the mean elevation determined immediately before the sequence was acquired and thus showing a horizontal plane without the mixing paddle. Afterwards, all datasets of all samples are multiplied by the same factor, so that the value range is between -1 and 1.

For numerical reasons, the reference values are also scaled to the interval [0,1]. In detail, this means that the reference points of the flow curve were divided by 300. For the regression parameters, the slope is not scaled and the y-axis intercept is divided by 200. Thus, the intercept point lies in a similar range of values as the slope. If two reference measurements are available, one of the two is randomly selected per training sample. In the validation and test, the mean value of the two measurements or regression parameters is chosen as the true value. Data augmentation was carried out by randomly changing the brightness and contrast of the orthophotos in a certain interval. The DEMs are augmented by adding the same random value to each pixel of the dataset.

To evaluate our results for each sample, the Mean Absolute Error (MAE) is used. This means that the absolute difference of

Table 1. Reference values determined with the rheometer. In addition, the parameters are given, which result from a regression between the support points 20 min^{-1} and 50 min^{-1} . In case two measurements were available per sample, the mean value is given. In addition, the difference between these measurements is shown in brackets (large differences are indicated in red). The sample number in the first column is only used to refer to individual samples.

Sample #	rheometer measurements									Regression param.	
	Torque [Nmm]									slope m [Nmm-min]	inters. n [Nmm]
	support points [min^{-1}]										
	2	4	6	8	10	20	30	40	50		
4	42 (2)	48 (0)	53 (2)	57 (-1)	59 (-2)	71 (-1)	78 (-1)	85 (0)	90 (-1)	0.6 (0.0)	58 (-1)
2	50 (-)	54 (-)	62 (-)	64 (-)	67 (-)	81 (-)	91 (-)	97 (-)	102 (-)	0.7 (-)	68 (-)
25	58 (-2)	64 (-1)	70 (2)	75 (-1)	79 (-1)	95 (4)	106 (3)	114 (3)	119 (-2)	0.8 (-0.2)	81 (8)
18	65 (-)	70 (-)	80 (-)	83 (-)	88 (-)	103 (-)	114 (-)	122 (-)	129 (-)	0.8 (-)	88 (-)
22	66 (2)	74 (0)	81 (0)	86 (0)	91 (-2)	105 (-2)	117 (-2)	127 (-1)	135 (-3)	1.0 (0.0)	87 (-1)
1	74 (9)	84 (8)	90 (2)	94 (-1)	98 (2)	120 (1)	130 (5)	141 (4)	149 (3)	1.0 (0.0)	101 (2)
21	76 (-1)	84 (-2)	92 (-2)	98 (1)	103 (-1)	121 (-2)	135 (-3)	145 (-2)	153 (-5)	1.1 (-0.1)	101 (-1)
16	77 (-)	86 (-)	93 (-)	98 (-)	105 (-)	121 (-)	132 (-)	131 (-)	146 (-)	0.8 (-)	107 (-)
20	78 (4)	86 (2)	93 (3)	100 (2)	105 (2)	123 (3)	137 (2)	145 (0)	154 (0)	1.0 (-0.1)	104 (6)
7	79 (-)	88 (-)	94 (-)	103 (-)	108 (-)	125 (-)	140 (-)	150 (-)	159 (-)	1.1 (-)	104 (-)
3	80 (-)	87 (-)	94 (-)	103 (-)	106 (-)	124 (-)	138 (-)	150 (-)	158 (-)	1.1 (-)	103 (-)
10	84 (-)	92 (-)	105 (-)	108 (-)	116 (-)	133 (-)	148 (-)	159 (-)	168 (-)	1.2 (-)	111 (-)
17	84 (3)	93 (1)	101 (1)	106 (4)	110 (3)	130 (-4)	146 (0)	155 (1)	163 (1)	1.1 (0.1)	111 (-6)
14	85 (1)	96 (1)	104 (2)	111 (2)	116 (2)	135 (3)	148 (2)	160 (4)	168 (4)	1.1 (0.0)	114 (1)
13	91 (3)	99 (3)	109 (1)	115 (0)	121 (-2)	139 (0)	154 (1)	166 (4)	175 (3)	1.2 (-0.2)	117 (2)
11	96 (-)	104 (-)	116 (-)	123 (-)	126 (-)	149 (-)	162 (-)	176 (-)	186 (-)	1.2 (-)	125 (-)
24	97 (2)	108 (4)	118 (3)	126 (3)	132 (2)	154 (-1)	169 (0)	183 (-2)	192 (-2)	1.3 (-0.1)	130 (0)
19	112 (6)	125 (7)	135 (10)	142 (7)	150 (9)	174 (8)	193 (11)	206 (13)	218 (12)	1.4 (0.1)	147 (6)
23	122 (-5)	136 (-1)	149 (1)	158 (2)	164 (-2)	191 (3)	209 (4)	224 (-2)	234 (-6)	1.4 (-0.3)	164 (11)
12	139 (2)	156 (2)	165 (1)	174 (2)	183 (4)	213 (0)	233 (6)	251 (3)	264 (5)	1.7 (0.1)	180 (-1)
9	148 (-)	166 (-)	180 (-)	190 (-)	202 (-)	234 (-)	259 (-)	278 (-)	293 (-)	1.9 (-)	198 (-)

each prediction to its true value is averaged over all samples. In addition, the standard deviation is used to show the precision of the predictions. For the experiments in which the 9 support points are predicted, the MAE is averaged over the support points for a more compact representation. Furthermore, the Root Mean Square (RMS) of the standard deviations (corresponding to the root of the mean variance) of the support points is calculated in these cases.

5.3 Results and Discussion

In this section, the results of the three experiments are presented and discussed.

5.3.1 Experiment series 1: In figure 6, the predictions for the images of the test samples of all test splits are plotted. In addition, the mean value over all predictions of a sample and the reference values are shown. Table 2 lists the MAE and the standard deviation across the predictions. The given standard deviation can also be seen visually in the plots. The MAE averaged over all support points lies between 3 and 10 for the *Default CNN* and between 3 and 14 for the *ResNet-18*. In general, it can thus be said that it is possible to predict the support points of the flow curves of individual samples with both architectures, while the *Default CNN* performs a little better in accuracy. For samples whose flow curves tend to be outside the range of the other values (samples 4 and 12), the prediction is not very accurate, however. Especially in split 3, an extrapolation has occurred in the test set for sample 4. The mean value of all predictions of sample 12 in split 2 agrees with the results of *Default CNN* (compare Fig. 6 (b)), but the standard deviation is rather high. With *ResNet-18* on the other hand, the predictions are more precise, but the mean value over all predictions is less accurate. The MAE is also lower for the *Default CNN* for sample 12.

If the differences between two measurements are looked at in table 1, it can be seen that also the determination in the

Table 2. Performance for *DefaultCNN* and *ResNet-18*: MAE and RMS of the standard deviation of the 9 predicted support points.

Sam. #	Split 1			Split 2			Split 3		
	17	23	18	11	3	12	11	4	21
MAE [Nmm]									
def. CNN	3	6	3	6	7	8	5	10	5
ResNet-18	3	9	3	5	4	13	6	14	6
RMS of standard deviation [Nmm]									
def. CNN	4	8	4	8	9	11	6	1	2
ResNet-18	4	8	4	6	5	6	6	3	5

viscometer is subject to errors. For example, the difference between two measurements in a support point for sample 23 is as high as 6. Under these circumstances, the prediction accuracy achieved can be considered a success. Overall, the *Default CNN* delivers a slightly lower MAE and is therefore also used for the further experiments. One reason for this may be that the much smaller number of parameters alone forces the network to generalise better. Moreover, the ResNet is pre-trained on RGB images, whereas here, the input is a greyscale image and a DEM.

5.3.2 Experiment series 2: Table 3 shows the results of the second series of experiments. It can be seen that the direct determination of the regression parameters does not have any advantages with respect to experiment 1, which is somewhat surprising. The reason may be that the regression parameters are somewhat prone to error, since they are only determined by four support points. For example, the difference in the slope of the regression lines from both measurements is 0.3, which is a large difference compared to the range of slope values of all samples lying between 0.6 and 1.9. If training is then carried out directly with such an erroneous measurements, this can lead to unstable results, whereas, when training is done with all 9 support points, a single inaccurate value is not as significant.

Table 3. Prediction of the regression parameters by (a) direct learning and (b) post-calculation from predicted support points.

	Split 1			Split 2			Split 3		
Sam. #	17	23	18	11	3	12	11	4	21
MAE slope m [Nmm-min]									
(a)	0.1	0.2	0.1	0.1	0.1	0.1	0	0.1	0.2
(b)	0.1	0.1	0	0.1	0.1	0.1	0	0.1	0.1
standard deviation slope m [Nmm-min]									
(a)	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0	0
(b)	0	0.1	0	0.1	0.1	0.1	0.1	0	0
MAE y-axis intercept n [Nmm]									
(a)	4	7	5	5	3	6	7	11	5
(b)	3	7	4	6	6	8	4	10	4
standard deviation y-axis intercept n [Nmm]									
(a)	5	7	6	6	4	6	6	4	4
(b)	4	7	4	8	8	10	5	1	2

5.3.3 Experiment series 3: Figure 7 and table 4 show the results of the third set of experiments. Similar observations can be made as for the previous experiments, such as the misclassification of sample 4. It should be noted that training only with orthophotos gives worse results than training only with DEMs. When using only images from sequences where the direction of movement of the mixing paddle is the same, the results show a higher accuracy for the mixing paddle moving away from the camera, suggesting that the groove the mixing paddle leaves behind is a clearer indication of the rheological properties than the somewhat elevated parts of the surface.

Table 4. Variation of input sources: MAE and RMS of the standard deviation of the 9 predicted support points.

	Split 1			Split 2			Split 3		
Sam. #	17	23	18	11	3	12	11	4	21
MAE [Nmm]									
def. CNN	3	6	3	6	7	8	5	10	5
DEM only	4	7	3	4	4	6	5	10	4
ortho only	5	23	6	6	11	21	7	15	16
away fr. cam.	2	4	3	5	3	3	4	9	4
tow. cam.	7	4	3	7	8	7	4	8	4
RMS of standard deviation [Nmm]									
def. CNN	4	8	4	8	9	11	6	1	2
DEM only	4	9	4	5	6	8	5	1	3
ortho only	7	9	6	8	9	16	9	4	5
away fr. cam.	3	3	4	4	4	5	5	1	2
tow. cam.	3	5	4	8	8	9	5	1	3

6. CONCLUSION AND FUTURE WORK

The results of this work have shown that it is possible to determine the rheological properties of ultrasonic gel, used here as a reference substance for cement paste, on the basis of image data, considering the potential inaccuracy in the determination of the reference values. This is in contrast to what experts can visually achieve: they typically can determine the approximate viscosity of concrete during the mixing process, but not the yield stress. It should be noted, that while viscosity and yield stress are actually two independent variables, in our experiments, these two quantities are highly correlated. Figure 5 shows this correlation. Thus, it is possible that in our experiments the yield stress (y-axis intercept) could only be determined by its correlation with the plastic viscosity (slope). This

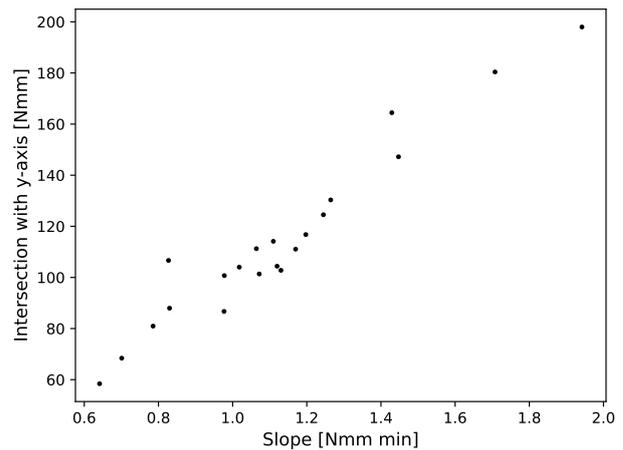


Figure 5. Correlation of the determined regression parameters. Each point represents one sample.

presumption needs to be tested in future research.

It is also noticeable that the standard deviation of the predictions of a sample can be relatively high, while still almost predicting the true value on average, leading to the assumption that random rather than systematic errors dominate the results. If this finding turns out to be correct, it should be considered for following applications to predict the rheological parameters based on several images where the mixing paddle is located at different positions in the image. In this sense, the use of the Long-Short-Term-Memory (LSTM) methodology (Hochreiter and Schmidhuber, 1997), which allows the processing of entire image sequences, should also be investigated. In this way, inaccurate results from some images could be compensated for.

In future work, the same setup will be tested with real concrete. Even though this work has shown that the prediction of the flow curve based on images only (here: orthophotos) is more difficult, monoscopic approaches should still be considered for use with concrete, because the appearance of concrete, such as colour and reflections, could be a reliable indication of its rheology.

ACKNOWLEDGEMENTS

This work is supported by the Federal Ministry of Education and Research of Germany (BMBF) as part of the research project ReCyControl [Project number 0336260A], <https://www.recycontrol.uni-hannover.de/de/>

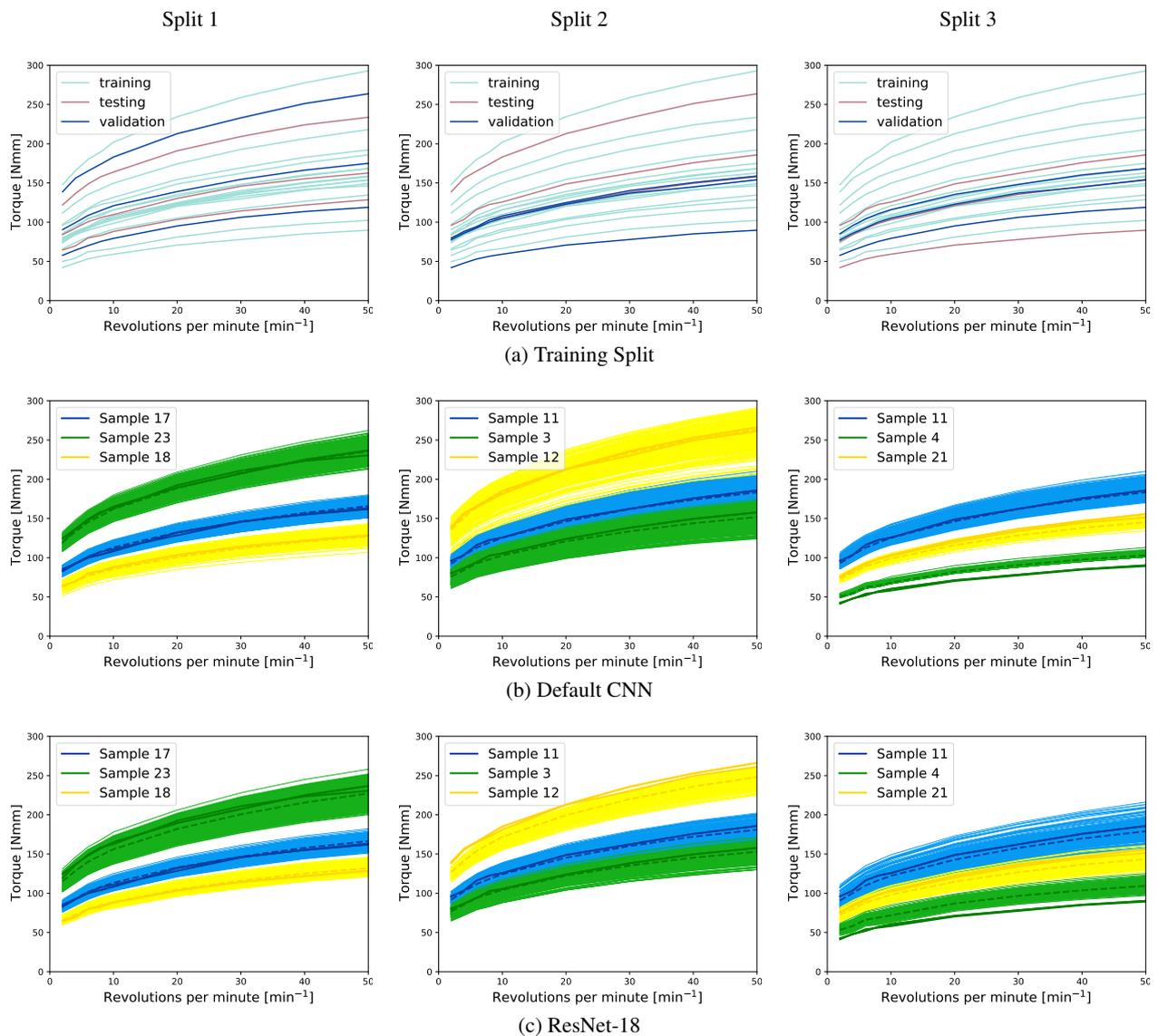


Figure 6. Results for the test sets of 3 different training splits. All predictions of the images belonging to a sample are shown with a thin line in the colour of the respective sample. The thicker dashed lines in the darker colours mark the average prediction values over all test images of a sample. The thicker solid lines in the same colour indicate the reference value (in case two measurements were available, both are plotted).

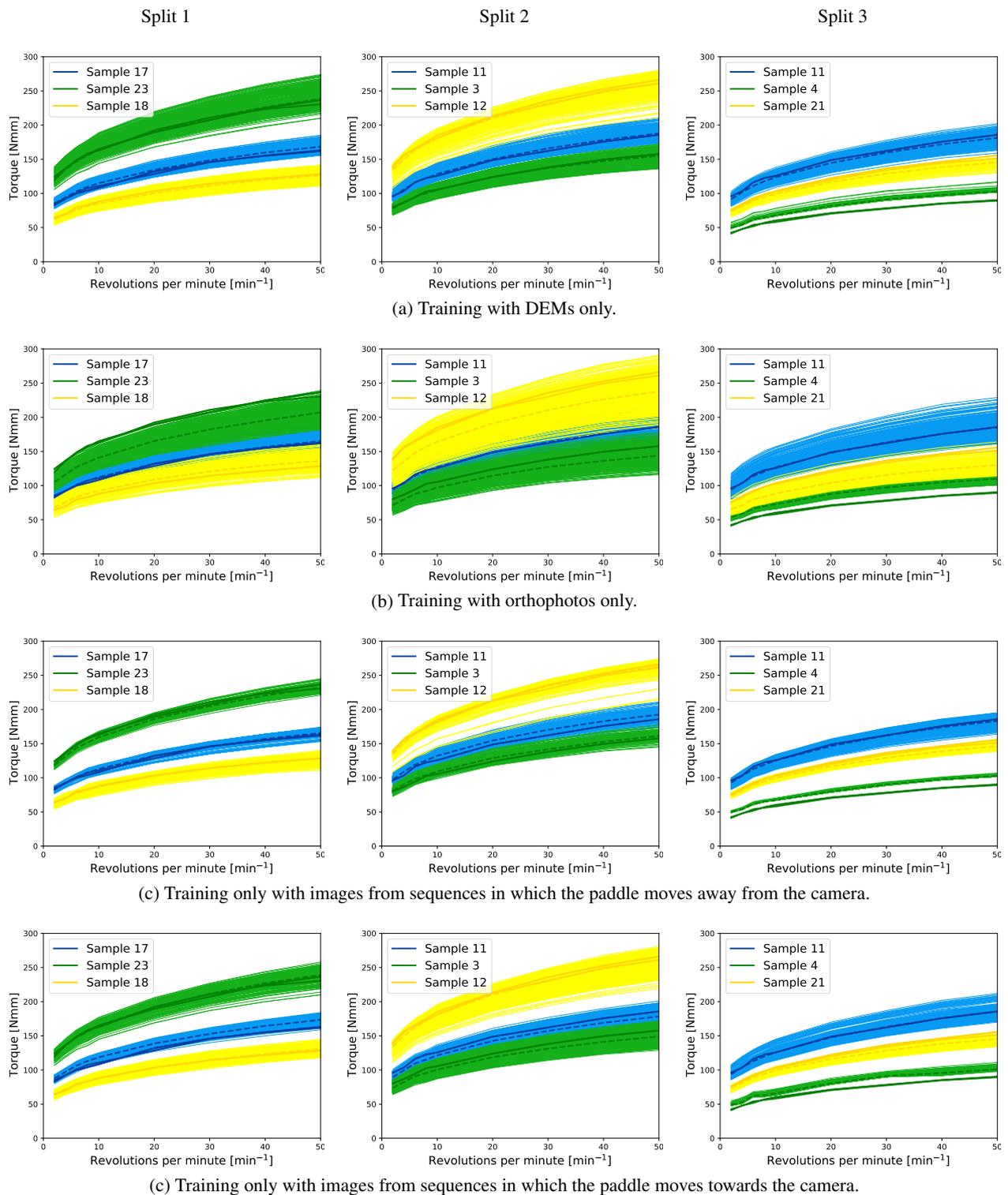


Figure 7. Results for the test sets of 3 different training splits. All predictions of the images belonging to a sample are shown with a thin line in the colour of the respective sample. The thicker dashed lines in the darker colours mark the average prediction values over all test images of a sample. The thicker solid lines in the same colour indicate the reference value (in case two measurements were available, both are plotted).

REFERENCES

- BTB, 2021. Jahresbericht 2021 des Bundesverbands der Deutschen Transportbetonindustrie e. V. (BTB). https://www.transportbeton.org/fileadmin/transportbeton-org/media/Verband/pdf/BTB_Jahresbericht_2021_online.pdf. (accessed 07/01/2022).
- Chidiac, S., Mahmoodzadeh, F., 2009. Plastic viscosity of fresh concrete – A critical review of predictions methods. *Cement and Concrete Composites*, 31(8), 535-544.
- Coenen, M., Schack, T., Beyer, D., Heipke, C., Haist, M., 2021. Semi-Supervised Segmentation of Concrete Aggregate Using Consensus Regularisation and Prior Guidance. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 52, 83–91.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database. *2009 IEEE conference on computer vision and pattern recognition*, 248–255.
- Ding, Z., An, X., 2018. Deep learning approach for estimating workability of self-compacting concrete from mixing image sequences. *Advances in Materials Science and Engineering*, 2018.
- Haist, M., Link, J., Nicia, D., Leinitz, S., Baumert, C., von Bronk, T., Cotardo, D., Pirharati, M. E., Fataei, S., Garrecht, H. et al., 2020. Interlaboratory study on rheological properties of cement pastes and reference substances: comparability of measurements performed with different rheometers and measurement geometries. *Materials and Structures*, 53(4), 1–26.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.
- Hochreiter, S., Schmidhuber, J., 1997. Long Short-term Memory. *Neural computation*, 9, 1735-80.
- Krizhevsky, A., Sutskever, I., Hinton, G. E., 2012. Imagenet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, NIPS'12*, Curran Associates Inc., Red Hook, NY, USA, 1097–1105.
- Lau Hiu Hoong, J. D., Lux, J., Mahieux, P.-Y., Turcry, P., Ait-Mokhtar, A., 2020. Determination of the composition of recycled aggregates using a deep learning-based image analysis. *Automation in Construction*, 116, 103204.
- Li, S., An, X., 2014. Method for estimating workability of self-compacting concrete using mixing process images. *Computers and Concrete*, 13, 781-798.
- Nerorth, G., Vollenschaar, D., 2011. Beton. *Wendehorst Baustoffkunde: Grundlagen – Baustoffe – Oberflächenschutz*, Vieweg+Teubner, Wiesbaden, 247–381.
- Nguyen, T.-D., Tran, T.-H., Hoang, N.-D., 2020. Prediction of interface yield stress and plastic viscosity of fresh concrete using a hybrid machine learning approach. *Advanced Engineering Informatics*, 44, 101057.
- Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition.
- Song, Y., Huang, Z., Shen, C., Shi, H., Lange, D. A., 2020. Deep learning-based automated image segmentation for concrete petrographic analysis. *Cement and Concrete Research*, 135, 106118.
- Statistisches Bundesamt (Destatis), 2021. Abfallbilanz (Abfallaufkommen/-verbleib, Abfallintensität, Abfallaufkommen nach Wirtschaftszweigen) 2019. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Umwelt/Abfallwirtschaft/Publikationen/Downloads-Abfallwirtschaft/abfallbilanz-pdf-5321001.pdf?__blob=publicationFile. (accessed 10/03/2022).
- Sutskever, I., Martens, J., Dahl, G., Hinton, G., 2013. On the importance of initialization and momentum in deep learning. *International conference on machine learning*, PMLR, 1139–1147.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1–9.
- Tuan, N. M., Hau, Q. V., Chin, S., Park, S., 2021. In-situ concrete slump test incorporating deep learning and stereo vision. *Automation in Construction*, 121, 103432.
- Yahia, A., Mantellato, S., Flatt, R., 2016. 7 - concrete rheology: A basis for understanding chemical admixtures. P.-C. Aïtcin, R. J. Flatt (eds), *Science and Technology of Concrete Admixtures*, Woodhead Publishing, 97–127.
- Yang, H., Jiao, S.-J., Yin, F.-D., 2020. Multilabel Image Classification Based Fresh Concrete Mix Proportion Monitoring Using Improved Convolutional Neural Network. *Sensors*, 20(16). <https://www.mdpi.com/1424-8220/20/16/4638>.