Classification and estimation of the mass composition of recycled aggregates by deep neural networks

Jerome Lux, Jean David Lau Hiu Hoong, Pierre-Yves Mahieux, Philippe Turcry

Classification and estimation of the mass composition of recycled aggregates by deep neural networks

Authors

Jérôme LUX, Jean David LAU HIU HOONG, Pierre-Yves MAHIEUX, Philippe TURCRY

Laboratoire des Sciences de l'Ingénieur pour l'Environnement (LaSIE), UMR CNRS 7356,

La Rochelle Université, Avenue Michel Crépeau, 17000 La Rochelle, France

Abstract

The construction sector is one of the largest consumers of natural resources, but also a producer of a considerable amount of waste. Construction and Demolition Waste can be transformed into recycled aggregates and used as a substitute for natural aggregates, either in road construction or concrete, which is one way to reduce the environmental impacts of the construction industry. In order to increase the use of recycled aggregates in high value-added materials like concrete, it is important to guarantee the quality of the produced aggregates. The recycling industry therefore needs new methods to automate the characterization of recycled aggregates.

In a previous work we showed that deep convolutional networks can classify the different constituents of recycled aggregates, achieving an average accuracy of 97%. In this work, we propose a novel network architecture called RACNET designed to

estimate the mass, the class and the binary mask of recycled aggregates from only 2D images. This could replace advantageously manual sorting tests as well as other geometric characterization tests (like particles size distribution), allowing for a real-time monitoring of recycled aggregates composition. We also present a prototype which preform automatic characterization of a flow of aggregates in real conditions, showing that our approach could be used in real industrial environments

Keywords: recycled aggregates; convolutional neural networks; deep learning; image classification; mass evaluation, circular economy

1 Introduction

Based on a report from the World Bank [1], approximately 1.3 billion tons of solid waste were produced each year on the planet in 2012, and this amount was forecasted to reach 2.2 billion tons in 2025 at latest. According to Transparency Market Research [2], half of the global amount of solid waste produced each year comes from construction materials. In the European Union, construction and demolition waste (CDW) accounted for 36% of the 2 337 million tons of waste produced in 2018 [3]. In France, the construction sector is the largest producer of waste and consumer of natural aggregates with 224 Mt of CDW, 69% of the total amount of waste, produced in 2017 [4] and 311 Mt of natural aggregates extracted the same year [5].

The recycling of CDW is an effective way to reduce their storage in landfills or in specialized sites and to decrease the extraction of natural aggregates at the same time. Consequently, the European Union and the French government have passed laws in order to promote circular economy in the construction sector through the recycling of

CDW. The European Union set up the Horizon 2020 program [6] to encourage its member states to recycle at least 70% of their inert and non-hazardous CDW.

Before being recycled, inert CDW are crushed and screened on recycling platforms in order to obtain recycled aggregates (RA). They are a complex mixture of different materials such as recycled concrete aggregates, natural stones, clay bricks, bituminous grains, along with traces of glass, wood and plastics. The nature of the materials that arrive on recycling platforms depends on the original site. Moreover, the crushing technique used to produce the RA affects the proportion of attached mortar on recycled concrete aggregates, as well as the amount of fines generated [7,8]. Hence, the composition of RA is variable.

Several studies have been carried out in order to investigate the effects of partial or total replacement of natural aggregates by RA in concrete manufacturing. Generally, this substitution degrades the mechanical properties and durability of concrete [9–11]. It has however been found that carbonation of recycled concrete aggregates is a solution to alleviate these issues [12–14]. This would also contribute to reducing the carbon footprint of the construction sector. Among the different materials present in RA, only recycled concrete aggregates can carbonate, thanks to their attached mortar. Thus, it's crucial to be able to determine the proportion of recycled concrete in a batch of recycled aggregates so that we can estimate the amount of CO_2 that may be potentially mineralized. In road construction, the use of RA containing a certain amount of bricks can even give better mechanical strengths than natural aggregates [15–18]. It is interesting to note that RA can also be used to make alternative hydraulic binders [19,20].

These studies show that the composition of RA influences their physico-chemical and mechanical properties and that it is essential to characterize them to be able to produce high grade materials. Currently, the EN 933-11 standard recommends manual sorting in order to determine that composition according to 6 classes: Rc (concrete), Rb (clay bricks or tiles, other earthenware or ceramic tiles), Ru (natural stones), Ra (bituminous materials), Rg (glass) and X (other). However, this procedure has some significant drawbacks.

Manual sorting is a time-consuming and tedious task which is carried out only occasionally on recycling platforms. Furthermore, since it is usually performed only once by a single operator on a particular sample of RA, the repeatability and reproducibility of the results are not evaluated. Especially, it is somewhat difficult to clearly differentiate recycled concrete aggregates from natural stones. The latter being used as aggregates in concrete production, some mortar can remain attached on them. According to the EN 933-11 standard, all concrete products or mortar should be considered as recycled concrete aggregates. This is arguable since a lot of them are mostly made of a natural stone with only a small part of attached mortar. Therefore, the properties of such aggregates would also be closer to those of the natural stone. This dilemma is also a source of confusion during manual sorting. Hence, it is very likely that there would be some variations in the results of the manual sorting if a particular sample of RA was analyzed several times by an operator or by different operators.

In a previous work [21], we showed that deep convolutional neural networks (CNN) can be used to predict the class of RA in an automated way. CNNs are networks that can be trained to perform certain tasks on digitized image, for exemple image classification which is the identification of the category to which an object belongs.

Concerning the classification of RA, the CNNs would be trained to identify the nature of the grain using their images. A labelled database consisting of thousands of images of individual RA, which have been previously sorted, had to be built in order to train the CNNs. The trained networks achieved a classification accuracy of 97% [21].

Since the composition of RA determined by the EN 933-11 is given in terms of mass proportions, we also estimated the mass of the grains that have been classified by the CNN using a simple geometrical approach based on a coefficient accounting for the shape and density of each RA class. This method was compared to manual sorting on test batch of RA and the results were very satisfactory, as there was less than 6% difference in mass proportions between the two methods. Most importantly, our method is much faster than manual sorting. Indeed, once the pictures of RA have been taken, it took less than five minutes to classify 3kg of RA using the CNN and to estimate their mass, while the manual sorting lasted for a few hours.

Nonetheless, the proposed method described above has some limitations. Indeed, the coefficients used to estimate the mass of the grains relied on the hypothesis that in a given sub-class, all the grains have the same shape and density. However, this is not always true as these parameters can be strongly affected by the crushing technique used [7,8].

In recent works, neural networks were successfully used to predict the mass of either individual or of sets of objects using 2D images. A first approach consists in measuring various geometrical features on the image using classical image analysis tools, which is then used as the input features of a Multi-Layer Perceptron (MLP). This approach was for example used in [22,23] in order to estimate the mass of fishes or ber fruits respectively. The limitation of this method is that it relies on a limited set of

arbitrarily chosen geometrical features. A more versatile and powerful approach is based on the use of Convolutional Neural Networks (CNN), which can learn features needed to correctly predict the mass using a 2D image as input. In [24,25], a regression head is attached to the top of a pre-trained classification CNN in order to predict the mass of objects. In these works, there is only one kind of material in the frames, so a global mass per frame can be computed without objects localization and classification. In [26], a more sophisticated architecture is proposed to estimate the mass of individual objects of different types. The proposed network is based on an Xception backbone and is made of two interconnected modules which predict the density and the volume of the object. This model makes use of another neural network that is pre-trained on the ShapeNet 3D object dataset to compute several geometrical features which are used in the two other modules.

All these works show the effectiveness of CNNs in predicting mass from 2D images. However, for this approach to be successful, one needs a very large dataset containing labelled objects as well as their ground truth mass to train the deep neural networks.

In this study, we describe a novel architecture of CNN designed to predict the class (i.e., the nature), the mass and the binary mask of recycled aggregates using only a 2D image (section 2.1). We have called this CNN the Recycled Aggregates Characterization Network (RACNET). In section 2.2, we describe the creation of the RA database, and we propose and validate a method to estimate the mass of each aggregates using careful sampling and shape factors. In section 3, results of classification, mass and area predictions are analyzed. We also give insights into what the neural network has learned. Lastly, we show in section 4 how the proposed method

139 can be fully automated so that it could be used in industrial applications using another
140 neural network to perform the automatic detection of aggregates.

141 2 Materials and methods

142 2.1 Recycled Aggregate Characterization Network (RACNET) architecture

143 Determining the mass of a 3D object using only 2D information can be divided
144 into two different tasks: estimating the object's density at first, followed by its volume,
145 as in [26].

146 The density is strongly correlated to the nature of the object, which can be
147 determined by a CNN as we showed in [21]. However, it can vary significantly in some
148 cases; for example, recycled concrete aggregates can be made of different types of
149 natural aggregates and the proportion of attached mortar can vary significantly from a
150 grain to another. The porosity, and therefore the bulk density, of some aggregates can
151 also vary significantly (e.g., limestone aggregates). The question is to what extent these
152 differences are visible in a picture.

153 The volume of an aggregate is related to its shape, which depends on both its
154 nature and the crushing process. Using only one camera, we have not access to any
155 volumetric information, but we can try to train the neural network to learn some 3D
156 features using the 2D picture. The volume can be expressed as the product between its
157 projected area A and an equivalent thickness T. The projected area is the area of the
158 mask of the aggregate in the 2D picture, which can be predicted by a trained CNN. The
159 equivalent thickness can be computed as the average of local thickness (i.e., by pixel).
160 Here, we cannot train the network to learn estimating the local thickness, as we do not

have access to this information, but this formulation guides us in deriving a relevant network architecture.

Following the previous discussion, the mass of an aggregate can be expressed as:

$$m = A \times T \times \rho \qquad \text{Eq. 1}$$

As said previously, we can measure the projected area A (in cm^2), but we have no information about the equivalent thickness T (in cm) and the bulk density ρ (in $\mathrm{g.cm}^{-3}$) of each aggregate. Note that we can only obtain average values of the densities from literature. Still, we follow this formulation and propose a new network architecture to estimate the mass of recycled aggregate as a product between three learned factors: the projected area A, a factor δ which is predicted on top of a trained classification network, and thus depends on features learned to classify the nature of the aggregate, and a second factor τ that is only slightly linked to the classification network and therefore depends on other meaningful spatial features.

The general architecture of this network called RACNET (**R**ecycled **A**ggregates **C**haracterization **NET**work) is depicted in Figure 1. It is based on a trained classification network, which can be any network like ResNet, ResNeXt, RegNet, EfficientNet [27–30], for example. In the proposed implementation, the backbone is a residual network using separable convolutions, as presented in [21]. We have found that this network performed better than other common architectures with comparable number of parameters, such as EfficientNetB3 or RegNetY1.6GF for the studied task. The backbone network is made of repetitions of a basic block called SepResBlock which is a residual block made up of two separable convolutions and a Squeeze-and-Excitation module (see [21,31] for more details). This network is made of 18 convolution layers and a dense head. Its number of parameters is about 11M.

After the final global average pooling, we attach a new head with two dense layers which combine previous features learned by the backbone network to classify the aggregates, in order to predict the δ factor. As this factor is computed by a combination of features used for the classification task, it will depend on the nature of the aggregate as well as other geometrical or textural characteristics specific to the RA class.

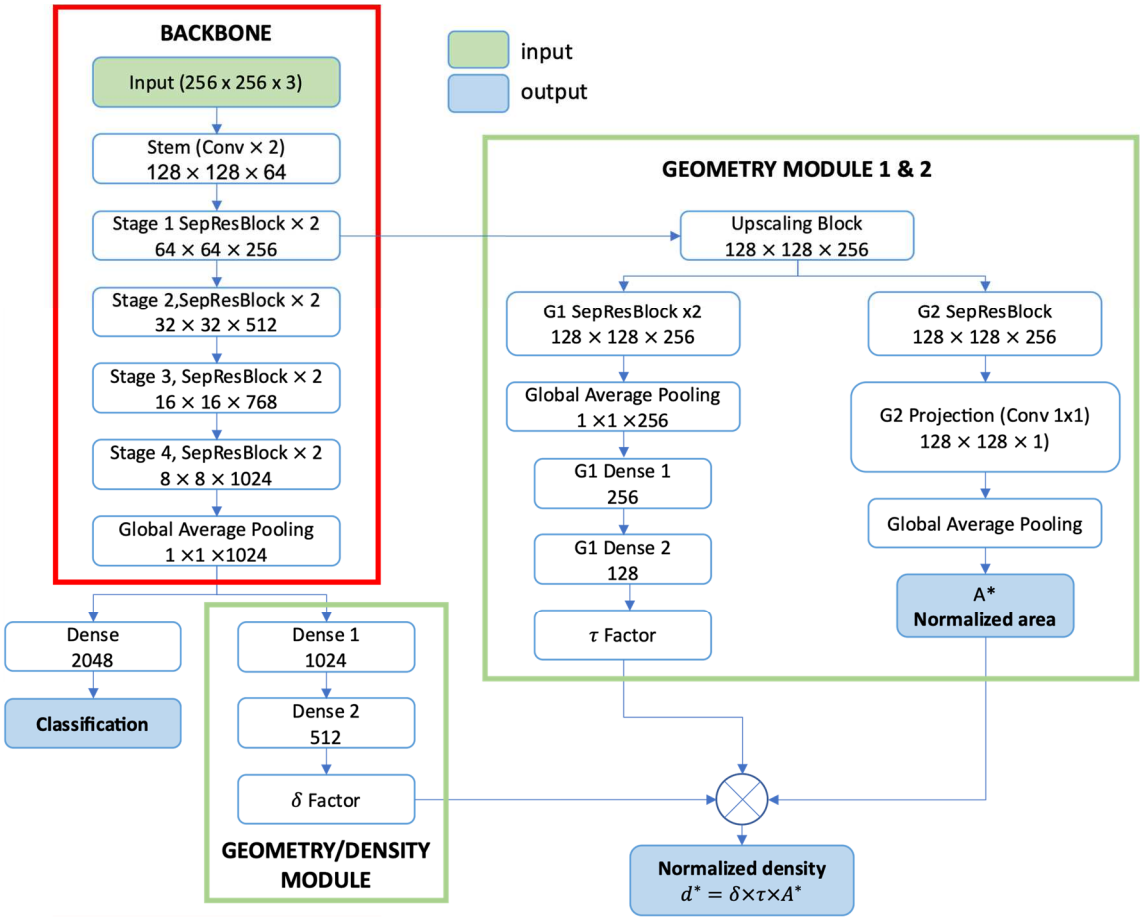


Figure 1: Proposed architecture of the RACNET. The stride s is given for each stage. Colored boxes indicate the input and targets used to train the network.

A new module called geometry module is attached to the output of the first stage. This module is therefore only slightly linked to the classification network and will learn whatever other useful features necessary to estimate the mass of aggregates. It is made of two complementary submodules G1 and G2. The first step consists in upscaling the feature map in order to have a sufficient spatial resolution. The first module G1 is made

of a repetition of residual convolutional blocks followed by a global average pooling layer connected to two dense layers. It outputs the single scalar τ, which will not necessarily depend on the class of the aggregate, but mainly on geometry features. The G2 module is responsible for predicting the binary mask of the aggregate, and it outputs the normalized area $A^* = A_d/(H \times W)$, where A_d is the discretized area, H and W are the height and width of the image respectively. This second module is made of a residual block followed by a projection and sigmoid activation, in order to get a feature map of dimensions $\frac{H}{2} \times \frac{W}{2} \times 1$. The normalized area is then obtained by a global average pooling of this last feature map.

The product of the three factors δ, τ and A^* gives a single scalar called normalized density and denoted d^*, which is related to the mass of the aggregates, as shown in the following.

$$d^* = \delta \times \tau \times A^* \qquad \text{Eq. 2}$$

The proposed implementation of the RACNET architecture has about 13,7 M trainable parameters, which makes it a very lightweight network.

2.1.1 Normalization of target values

The input of the convolutional neural network is a 2D image with fixed size $L \times L$ (in pixels). In this work, L is equal to 256 pixels. The predicted "mass" value should therefore be independent from the image true size. If we note V_d the discretized volume of the aggregate in voxels and R the resolution of the image in $\mathrm{pixel/cm}$, then the mass can be expressed as:

$$m = \rho \times \frac{V_d}{R^3} \qquad \text{Eq. 3}$$

We could use $m \times R^3 = \rho \times V_d$ as the target for the predicted mass value, but this product can take values from 1 to 1.3×10^4 g.voxel.mm^{-3}. To normalize the target, we simply divide it by a reference volume which is the volume of the cube with the same length as the image side L in pixel (here $L^3 = 256^3$).

The target d^* for the numerical mass output of the neural network is then defined as:

$$d^* = \frac{m \times R^3}{L^3} \quad \text{Eq. 4}$$

d^* is expressed in g/cm^3, and is therefore called a normalized density. We have tried other nonlinear normalization methods (such as quantile transform) but it was found that it hurts the learning process and decreases the final performance of the network. It seems important to preserve the relative differences between the target values, even if it gives some targets values as small as $6 \times 10^{-5} g/cm^3$ in our current database.

As stated previously, the area target is the area of the mask of the aggregate measured in number of pixels and denoted A_d, divided by the total image area:

$$A^* = \frac{A_d}{L^2} \quad \text{Eq. 5}$$

2.1.2 Loss

The RACNET predicts a one-hot encoded vector of class predictions, a normalized area and a normalized density. The total loss is expressed as the weighted sum of three losses:

$$L_{tot} = \alpha_1 L_{cls} + \alpha_2 L_{area} + \alpha_3 L_{density} \quad \text{Eq. 6}$$

Where L_{cls} is the categorical cross-entropy loss, and L_{area} and $L_{density}$ are the mean absolute percentage error (MAPE) of the predicted area and density respectively

238 and a_i are the corresponding loss weights. Due to the very wide range of values for
239 density targets, we have found that MAPE loss gives far better results than other losses
240 such as mean squared error (MSE).

241 **2.2 Recycled aggregates database**

242 The EN 933-11 standard gives 6 classes of RA, namely concrete grains (Rc),
243 natural stones (Ru), ceramics (Rb), bituminous grains (Ra), glass (Rg) and others (X).
244 However, some classes contain aggregates with considerable differences in properties
245 such as mechanical strength, density, water absorption, porosity and so on. For example,
246 limestone and basalt are both classified as natural stones but have very different
247 properties. Thus, we based ourselves on the 6 classes in order to define 16 sub-classes
248 regrouping RA with similar properties. For example, limestone and basalt are in two
249 distinct sub-classes (Ru01 and Ru02 respectively) that belong to the class natural stones
250 (Ru). The classes of the EN 933-11 standard were therefore divided into more
251 homogenous sub-classes as shown in Table 1 (see also [21]).

EN 933-11 classes	Corresponding sub-classes	Description
Rc	Rc	Concrete grains
Ru	Ru01	White stones such as limestone
	Ru02	Grey stones such as basalt & others of similar colors
	Ru03	Light colored grainy stones (majority of quartz and feldspar)
	Ru04	Colored or dark colored siliceous and rather angular stones
	Ru05	Light-colored and rounded alluvial stones
	Ru06	Slate
Rb	Rb01	Clay bricks or roof tiles
	Rb02	Ceramic tiles, earthenware tiles, etc.
Ra	Ra	Bituminous grains
Rg	Rg	Glass
X	X01	Wood
	X02	Plastics
	X03	Steel
	X04	Paper and cardboard
	X05	Others

Table 1: EN 933-11 classes and corresponding sub-classes of recycled aggregates

In the present work, we use two image databases. The creation of the first database, called D1, was detailed in a previous paper [21]. This database contains now about 38k images of aggregates of granulometric fraction 4/31.5 divided in 16 classes. The pictures are taken on a copy stand with a blue background at different resolutions, namely 23 pixels/mm for the 4/10 granular fraction and 12 pixels/mm for the 10/31.5 granular fraction. With this setting, the smallest diameter of an aggregate is discretized by a least 100 pixels approximately.

The second database, called D2, is made of images of RA of granulometric fraction 4/40 photographed with a linear camera on a conveyor belt. The experimental

setup is shown in Figure 2. It is made of a vibratory feeder which uniformly distributes the aggregates on a conveyor belt. Note that it is important to adjust the vibratory feeder to limit the overlapping of aggregates. Images of the aggregates are captured by a linear camera of 8192 pixels width. In the following, we work with images of 8192×4096 pixels (see Figure 3). The vertical position of the camera is set in order to obtain a resolution of about 28.4 pixels/mm, which allows to capture fine details on the smallest aggregates (down to 4 mm). Using this device, we took about 540 pictures for a total of 20k aggregates (see an example of a picture in Figure 3).

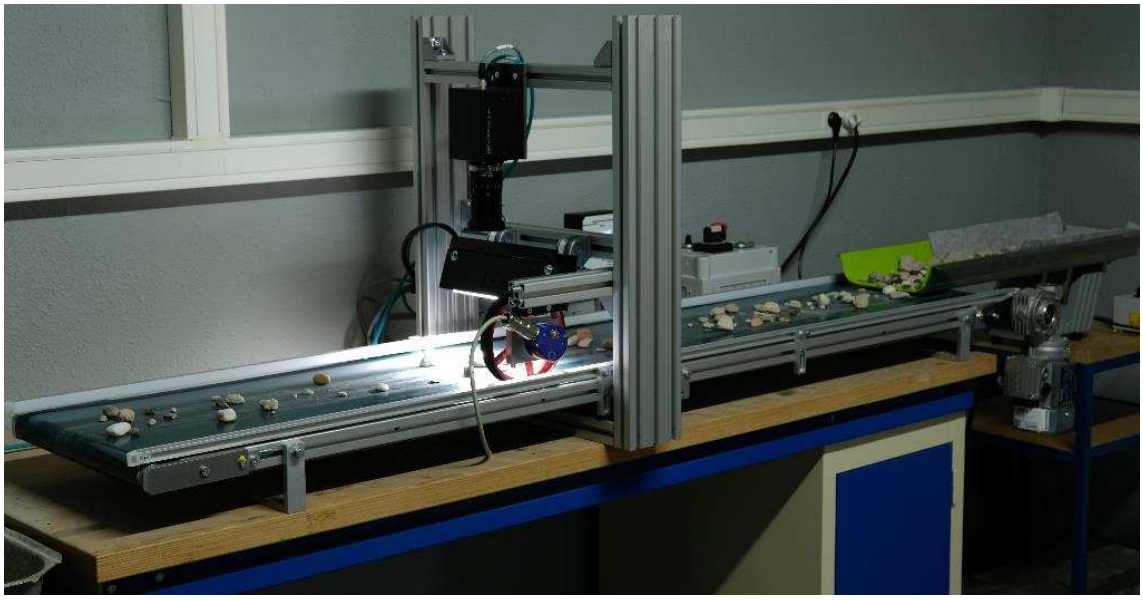


Figure 2 Automatic aggregates characterization device. A vibratory feeder (right) distributes aggregates on the conveyor belt. A linear camera capture pictures of the aggregates.

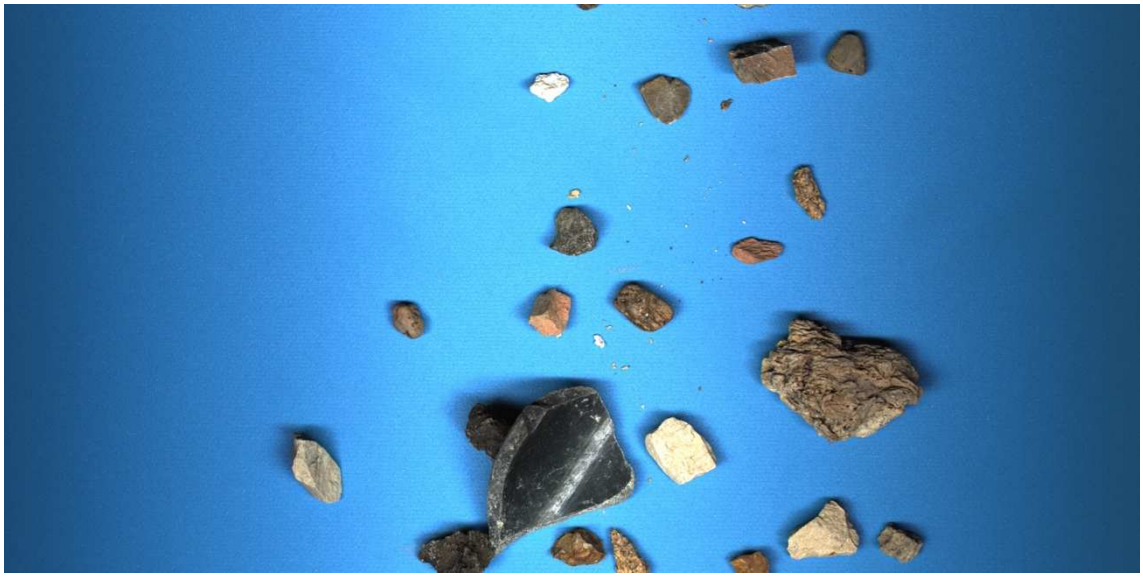


Figure 3: Mixed aggregates photographed by the linear camera (4096×8192 pixels)

Details about the number of elements per class are given in Table 2. Note that the databases are very unbalanced, as some classes are hard to find in industrial recycled aggregates (mostly X and Rg). These classes (greyed out in Table 2) are not used for evaluating the performance of the RACNET architecture.

Class	D1	D2	TOTAL
Ra	5659	329	5988
Rb01	3140	1406	4546
Rb02	2344	206	2550
Rc	7513	9396	16909
Rg	43	10	53
Ru01	6525	3241	9766
Ru02	3322	2890	6212
Ru03	1203	35	1238
Ru04	2947	921	3868
Ru05	3414	1008	4422
Ru06	1075	29	1104
X01	926	93	1019
X02	475	12	487
X03	281	25	306
X04	93	3	96
TOTAL	**38960**	**19604**	**58564**

Table 2: Number of elements per class in database D1 and D2. Greyed lines correspond to classes that are not used to train and evaluate the performance of RACNET.

2.3 Ground truth mass estimation

It is obviously not possible to weigh each aggregate individually, as it is very time consuming. We have chosen to weigh only a small part of the aggregates (more than 1800, from 8 different classes) to get sufficient ground truth mass values. For the other aggregates, we use a shape factor approach [21] to estimate the mass of individual grains knowing the total mass of the aggregates on each picture or series of pictures. The equivalent thickness in Eq. 1 can be expressed as the product of a characteristic thickness with a shape factor F. In this work, the characteristic thickness of an aggregate is defined as the diameter of the largest inscribed circle in the projected area of the aggregate as in [21]. For this approach to be successful, the aggregates on a picture must be of the same class and have similar morphological characteristics. We therefore proceed as follows:

- After manual sorting of a batch of aggregates coming from the same test batch, we weigh a small sample of aggregates of the same class and granular fraction.
- We take a picture or a series of pictures of this sample.
- Each aggregate is segmented individually using either classical image analysis tools or a home-made instance segmentation network, as it will be detailed later in section 4.
- Knowing the total mass M_k of the aggregates in a picture k (or a series of pictures), the mass of a single grain m_i is estimated as:

$$m_i = \rho_k F_k A_i D_i \qquad \text{Eq. 7}$$

Where A_i is the projected area (i.e., the area of the segmented mask), D_i the diameter of the largest inscribed circle in the grain and $\rho_k F_k$ the shape factor-density product for the picture k, defined as:

$$\rho_k F_k = \frac{M_k}{\sum_i A_i D_i} \qquad \text{Eq. 8}$$

With this method, we ensure that the sum of the estimated mass for all aggregates in a picture or a series of picture is equal to the real measured mass. Using careful sampling, we hope to limit the error on individual mass estimation.

In order to validate the proposed approach to estimate the mass of a single aggregate using the total mass of the aggregates in a picture or a series of picture, we compare the real mass of about 1800 weighed aggregates to the estimated mass calculated using shape factors. Figure 4 shows a scatter plot of the estimated mass versus the ground truth mass of each aggregate. Qualitatively, these results indicate that the procedure based on shape factors allows to estimate correctly the major part of the aggregates masses. Only a few lightweight aggregates show large relative errors. The mean absolute error for each class is presented in Table 3. The mape is quite similar for

all classes, between 20 and 30%, except for Ra, where it is only 16%. It indicates that the proposed performs well for all tested classes. If we plot the Absolute Percentage Error (APE) distribution (Figure 5), we can see that the APE is lower than 25% for about 61% of the aggregates and it is lower than 50% for 91% of the aggregates.

Using the proposed shape factor procedure along with careful sampling is therefore a good solution to estimate the ground truth mass of each aggregate when only the total mass of a batch of aggregates is available.

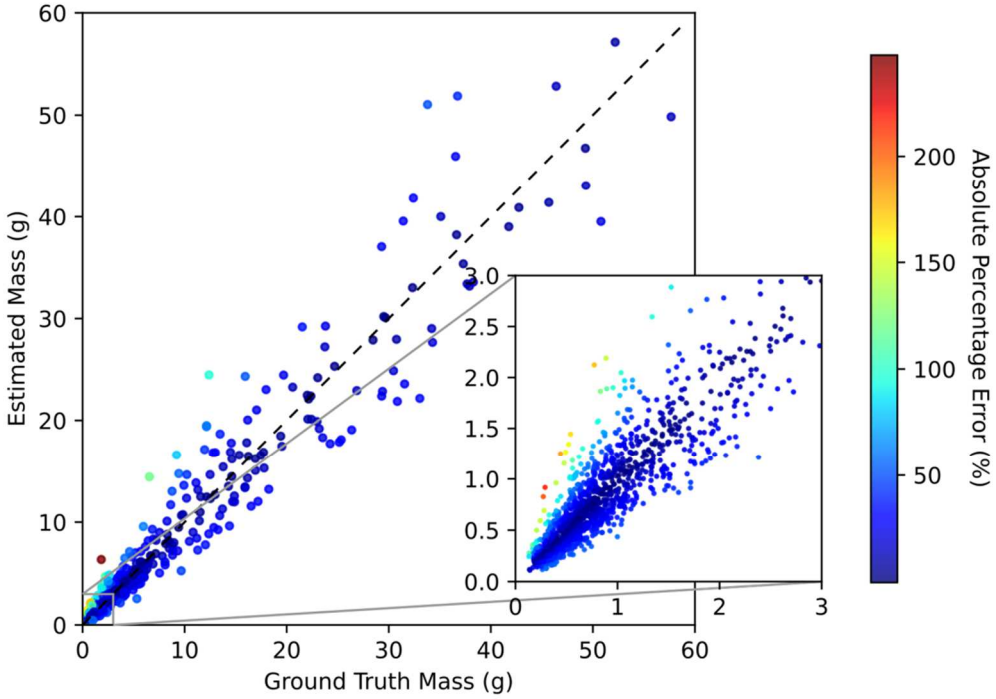


Figure 4 : Estimated mass vs ground truth mass. Markers are colored by the value of the Absolute Percentage Error. The dashed line is the identity line. A zoomed-in inset is plotted to see the details for lightweight elements.

Class	MAPE (%)
Ra	16.4
Rb01	28.5
Rb02	21.1
Rc	20.1
Ru01	21.0
Ru02	27.9
Ru04	29.0
Ru05	27.6

Table 3: mean absolute percentage error per class between ground truth and estimated mass using shape factor approach.

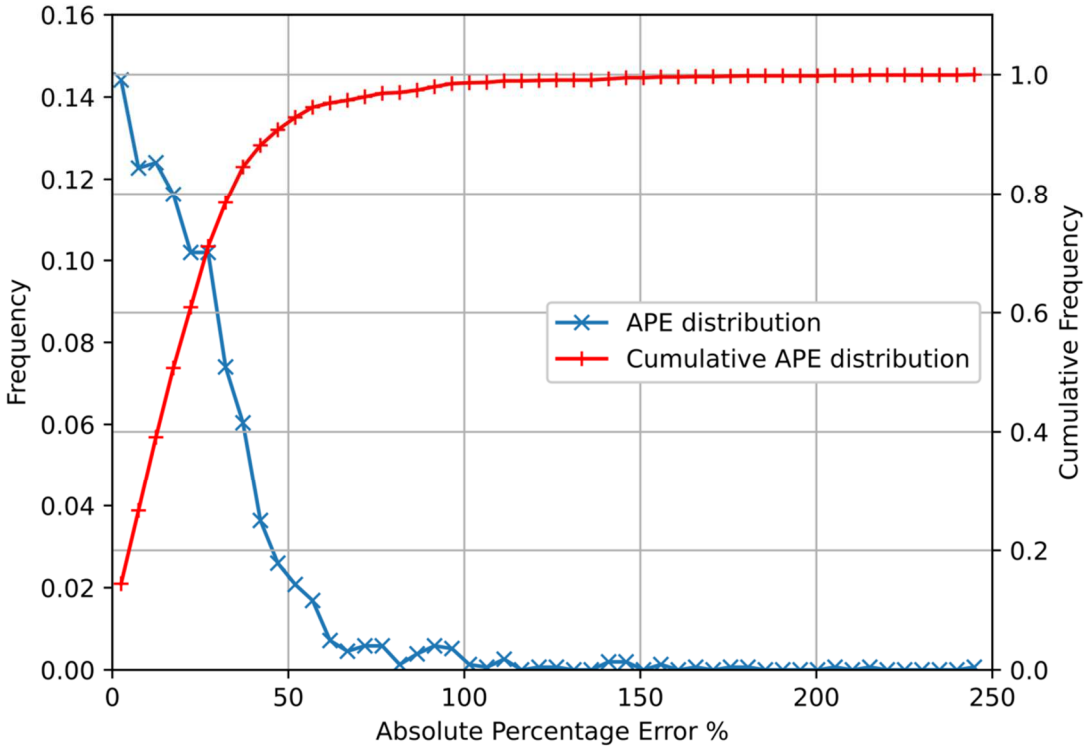


Figure 5 Distribution of Absolute Percentage Error between true and estimated mass.

3 Results and discussion

3.1 Training

The backbone and the classification head are trained beforehand. The training, validation and tests sets are generated using a mix of the D1 and D2 datasets. The target numbers of images per class in training/validation/test sets are 8000/800/800 (4000/400/400 from each dataset). If the number of elements is not sufficient, an image can be used twice using image augmentations (random rotation, gamma correction and noise, see [21]). Of course, even with augmentations, some classes do not contain enough elements to obtain the prescribed number of images. In this case, we try to keep the same split between the three sets. Finally, we discard classes with too few elements (Rg, X02, X03 and X04), keeping only 11 classes, as said previously. The composition of the training/validation/test dataset is given in Table 4.

Class	Training	Validation	Test
Ra	4544	454	427
Rb01	6336	634	517
Rb02	4250	424	409
Rc	8000	800	800
Ru01	8000	800	670
Ru02	8000	800	640
Ru03	2062	206	205
Ru05	5674	568	483
Ru04	5528	552	477
Ru06	1838	184	181
X01	1694	170	161
TOTAL	**55926**	**5592**	**4970**

Table 4: Number of elements in training/validation/test sets.

All the images are resized or padded to 256×256 pixels. In preliminary studies, we found that the classification accuracy depends on the image resolution. The size of 256×256 pixels was chosen because it was a good compromise between computation

speed/memory footprint and precision. A larger size did not provide significant improvement.

The classification network was trained using an Adam optimizer with an initial learning rate of 1.10^{-3} and a batch size of 32. The learning rate is reduced by a factor 2 when the loss does not decrease for 5 consecutive epochs. The training is stopped when the validation loss does not decrease during 25 consecutive epochs.

The 3 RACNET modules were then trained using a frozen backbone and with the same dataset and learning procedure, but with a batch size of 64. Mass and area targets were normalized as explained previously. The losses weights α_2 and α_3 were set to 1. Note that a dropout rate of 0.4 is applied during training between each dense layer in the DENSITY HEAD and the G1 GEOMETRY module.

The training and validation loss curves for normalized density and area are plotted in Figure 6. The overfitting is very limited, probably thanks to the dropout layers. The model used for prediction hereafter is the one with the lowest total validation loss (Eq. 6 using $\alpha_i = 1, i = 1,2,3$).

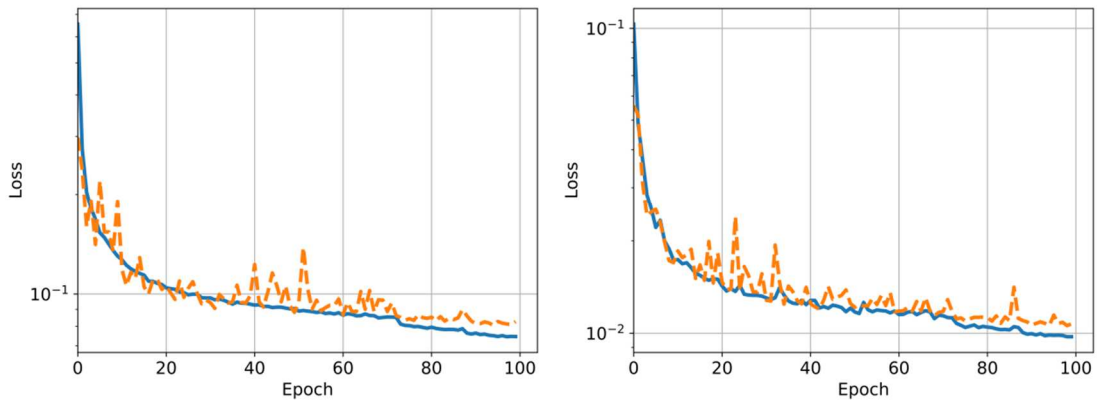


Figure 6 Training (solid blue line) and validation (dotted orange line) loss for normalized density target (left) and area target (right).

3.2 Classification results

Results presented hereafter are obtained on the test set containing about 4970 aggregates for a mass of about 6,14 kg.

The backbone achieves an average accuracy on the test set of about 98% which is similar to previous results using a network with nearly double number of parameters [21]. The fact that this lightweight backbone with only 18 convolution layers performs as well as a larger network may be due to the increase of the number of images in the training set and to a slightly more aggressive strategy for image augmentation (mostly the random addition of different types of noise on the aggregates). Complete classification results are summarized in the confusion matrix (Figure 7). The performance of our network is very satisfying, as the accuracy is greater than 95% for all classes. Most confusions occur between concrete (Rc) and natural stones (limestone Ru01 or basalt Ru02). This is mainly due to the fact that recycled concrete aggregates often contain natural stones with barely any attached mortar. Note that this is an inherent problem of the classification proposed in the EN 933-11 standard. A possible solution to detect correctly the attached mortar could be to segment mortar pixels using a specific semantic segmentation module.

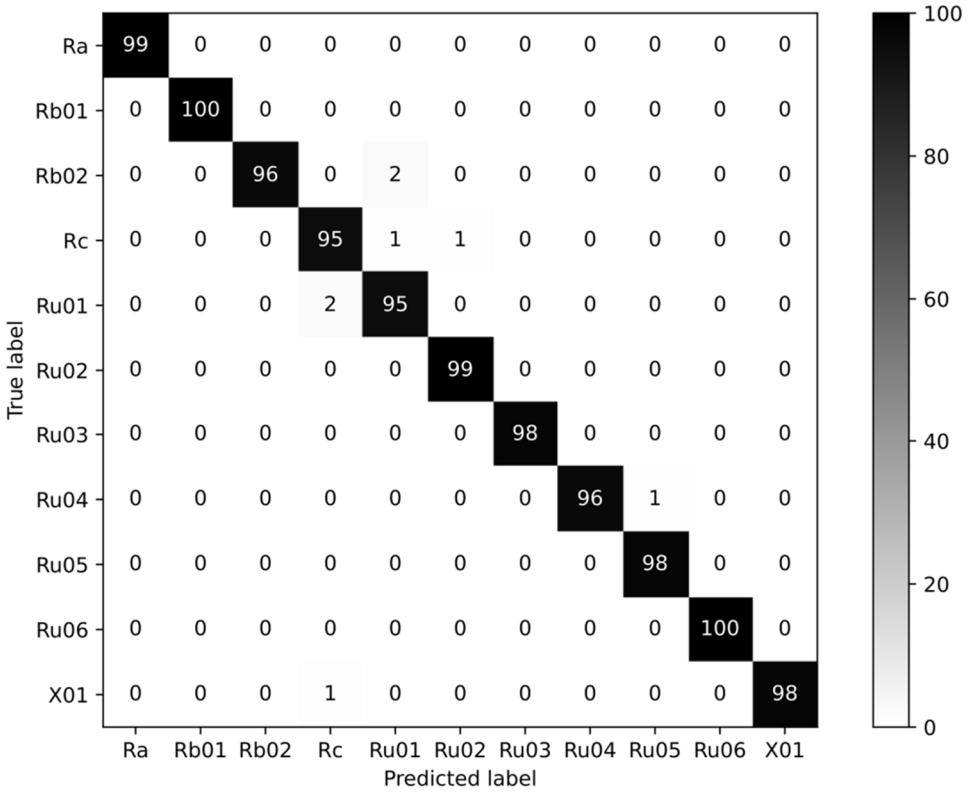


Figure 7 Confusion matrix for the RACNET classification backbone.

3.3 Mass regression

In order to evaluate the ability for the RACNET to correctly predict the mass of recycled aggregates, we can compute the Mean Absolute Percentage Error (MAPE) for each class. Although the MAPE is a good indication of the network performance, it does not reflect the fact that a large relative error is less important on a light aggregate than on a heavy one. From an industrial point of view, it is therefore more useful to quantify the relative error on the *total mass* for each class, because the most relevant feature is the precision of the global mass estimation for each class. We define the total mass absolute percentage error for a class k as the absolute relative error between the total ground truth mass M_{GT}^{k} and the total predicted mass M_{PRED}^{k} in the class k:

$$\mathrm{TM\_APE_k} = \frac{\left|M_{GT}^{k} - M_{PRED}^{k}\right|}{M_{GT}^{k}} \qquad \text{Eq. 9}$$

The MAPE and TM_APE for each class are summarized in Table 5. The total ground truth and predicted masses are also given for each class, as well as for the entire test set. The MAPE is lower or equal to about 10% for all classes except for wood aggregates (X01). This can be due to the low resolution of lots of the wood aggregates. Indeed, as the wood aggregates are often very elongated, the downscaling factor to resize the image to the chosen fixed input size can be larger than that of more circular aggregates. Moreover, the shape and real density variability is more important for this kind of aggregates. The MAPE is similar for the other classes, but we can notice that the error is more important for the classes containing a greater diversity of aggregates such as Ru04 or Rb02.

The global mass estimation for each class is quite good, as the average error is of only 2% and the relative errors are below 5% except for wood aggregates (7%). This result is very encouraging.

In Table 6, we also provide the comparison between the true and predicted mass fraction distribution (respectively denoted %GT and %PRED). It shows that the difference between the ground truth mass fractions and those predicted by the RACNET are all lower than 1 point.

These good results show that our network could be used as a replacement for manual sorting in industrial applications.

Class	M_{GT} **(g)**	M_{PRED} **(g)**	**MAPE (%)**	**TM_APE (%)**
Ra	383.1	393	5.68	2.56
Rb01	514.1	509	6.93	1.00

Rb02	891	885.7	8.64	0.60
Rc	933.4	901.4	7.87	3.43
Ru01	946.2	939.6	7.24	0.70
Ru02	417.4	412.5	6.30	1.17
Ru03	159.8	155.3	5.93	2.80
Ru04	527.1	533.1	8.50	1.13
Ru05	886	883.5	6.84	0.29
Ru06	56.7	57.5	6.05	1.36
X01	425.7	457.1	23.23	7.37
Average			**8.47**	**2.04**
Total	**6141**	**6127**		**0.21**

Table 5: total ground truth and predicted mass per class, as well as mass MAPE and relative error on total mass per class.

Class	**%GT**	**%PRED**
Ra	6.2	6.4
Rb01	8.4	8.3
Rb02	14.5	14.5
Rc	15.2	14.7
Ru01	15.4	15.3
Ru02	6.8	6.7
Ru03	2.6	2.5
Ru04	8.6	8.7
Ru05	14.4	14.4
Ru06	0.9	0.9
X01	6.9	7.5
Total	**100**	**100**

Table 6: Ground truth and predicted mass fraction for each class.

3.4 Segmentation masks

While the normalized area A^* is a target value used during training, there is no explicit loss for the mask. We can verify that the output of the G2 projection layer (after sigmoid activation and a threshold of 0.5) matches the mask of the aggregates by plotting the contour of the binary mask superimposed onto the input picture (Figure 8). The masks are qualitatively accurate, which is confirmed by the area MAPE loss which

is as low as 0.13% (see Figure 6). The binary mask can be used to measure some useful morphological characteristics (particle size distribution, angularity, etc.).

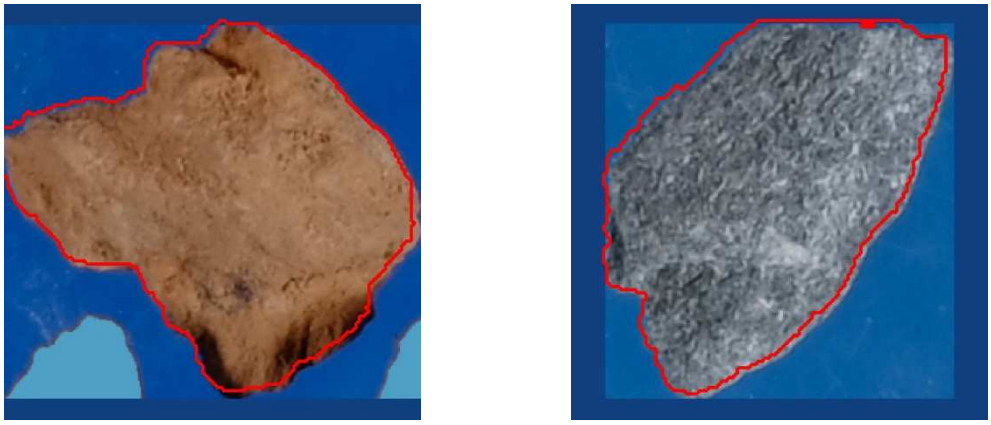


Figure 8 Contours of binary mask in red obtained after the G2 projection layer (with threshold 0.5).

3.5 What does the network learn?

The RACNET was designed so that the δ factor is strongly correlated to the class of the aggregates while the τ factor is only slightly linked to the classification network. As there are no target values for these two factors, we do not know *a priori* what the network will learn during training. From Eq. 4, we have:

$$d^* = \delta \times \tau \times A^* = \rho V \frac{R^3}{L^3} \qquad \text{Eq. 10}$$

where $V = \frac{V_d}{R^3} = A \times T$ is the real volume in cm^3, A the real area (in cm^2) and L the length of the image in pixels.

We know that $A = \frac{A^* L^2}{R^2}$, so the product $\delta \times \tau$ can be expressed as:

$$\delta \times \tau = \rho T \frac{R}{L} \qquad \text{Eq. 11}$$

In order to know what the RACNET actually learns, predicted values of δ and τ are plotted in Figure 9 for each class. We plot the normalized factor τ and not directly $T = \tau L/R$ because the equivalent thickness depends on the size of the aggregates,

445 while τ is independent of the resolution and size of the image and should depend on the
446 shape only. Here, we assume that δ is a density in g/cm^3. It is interesting to note that the
447 results tend to confirm our initial intuition. Indeed, δ factor values are in the same order
448 of magnitude as real aggregates density: above 2g/cm^3 for most natural stones and
449 concrete (Ru, Rc) and about 1.5g/cm^3 for clay bricks (Rb). Furthermore, relative
450 differences between the classes are quite realistic. Indeed, the predicted density is
451 smaller for more porous materials (like clay bricks or wood). One can however note that
452 the predicted δ factor is smaller than real average values for wood (0.28g/cm^3) and slate
453 (0.8g/cm^3). It is possible that some shape characteristics that are specific to all or most
454 elements of a class are integrated in the δ factor. Indeed, slate equivalent thickness is
455 roughly the same as other aggregates, whereas slate aggregates are flatter and thinner
456 than other aggregates. Regarding the predicted normalized equivalent thickness τ, the
457 values are similar for all classes except for wood aggregates. This may be due to the fact
458 that most of the wood aggregates are very elongated. In any case, this reflects the fact
459 that the shapes of wood aggregates are different from other aggregates. Note however
460 that there is no significative difference between the average normalized equivalent
461 thickness of rounded aggregates (Ru05) and other angular aggregates. Again, it may be
462 due to the fact that the roundness is a class characteristic and is integrated in the δ factor
463 rather than in the τ factor.

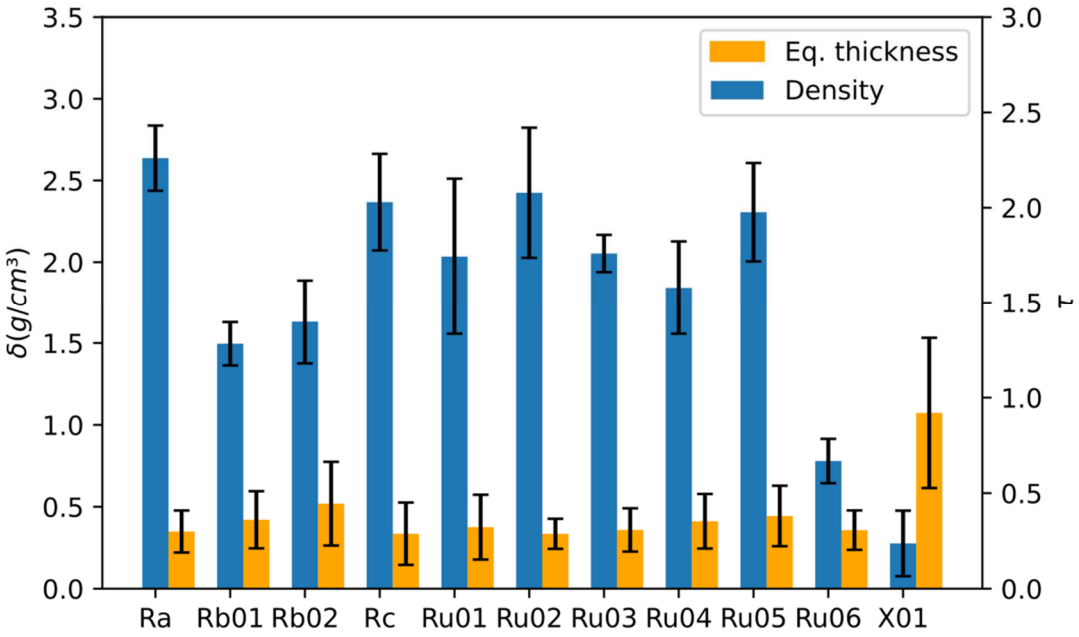


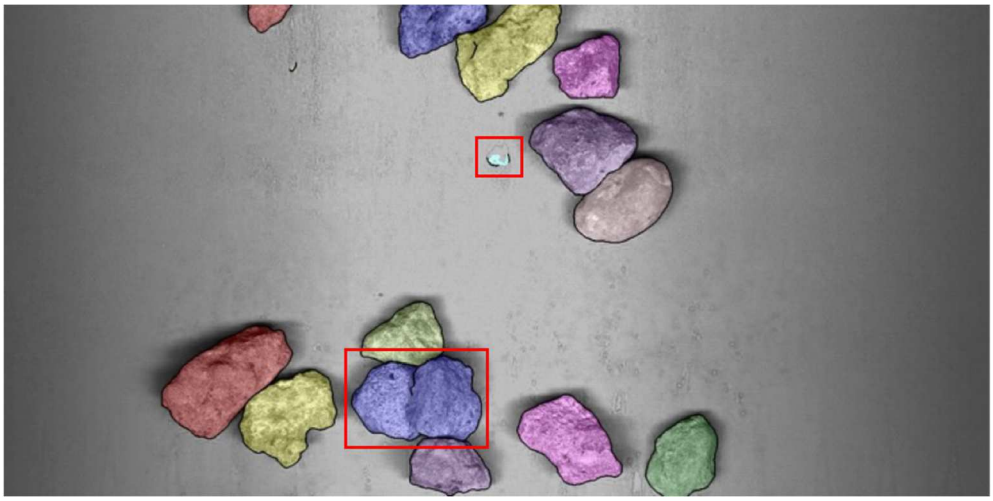
Figure 9 Histogram of predicted normalized density (δ factor) and equivalent thickness (τ factor) per class. Error bars represent standard deviations.

4 Automation and industrial application

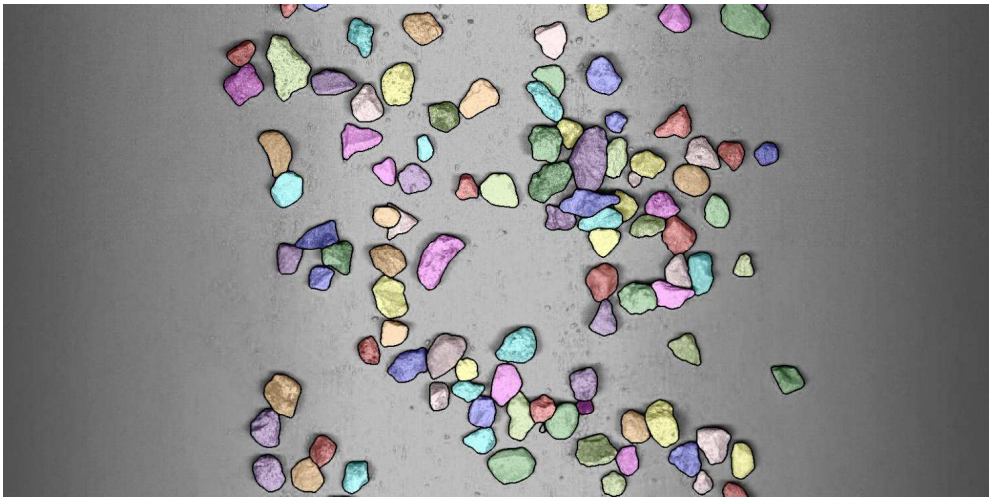
4.1 Extraction of individual aggregates

The main limitation of the RACNET is that the network's input must be a picture of a single aggregate. We therefore need to extract each aggregate individually, without its close neighbors. This task, called instance segmentation, is a very active research area. Architectures like MASK-RCNN [32] or more recently CondInst [33] or YOLACT++ [34] can be used to extract the bounding box, the class and the mask of each object in an image. Other approaches are based on edge detection [35–37]. In this work, we follow this path and we implemented a slightly modified version of ResUNet-a architecture [35]. This network outputs three feature maps: an object score map, a contour score map and a distance map. In order to get the instances, the contour map is first thresholded and dilated by a small factor, and subtracted from the binary object

map. A simple connected component labeling is then used to extract each aggregate. We also tried to use the distance map and a watershed algorithm, but it leads to over segmentation. The benefit of this approach is that the output of the network has the same resolution as the input image, giving far more precise masks than MASK-RCNN for example. It is also better suited than FCOS or MASK-RCNN networks to detect elongated and thin objects like wood particles or steel bars. The downside is that aggregates touching or slightly overlapping each other are sometimes extracted as a single instance (see Figure 10). This is not too detrimental because it often happens in the case of very similar aggregates, so it is not a problem for the classification task. Results also show that this does not reduce the accuracy of the mass prediction. Detailed architecture and performances of the detection network are given in APPENDIX.



(a)



(b)

Figure 10: Example of aggregate segmentation obtained after processing the ResUNet-a network outputs. In image (a), two overlapping aggregates (in the biggest red box) are detected as a single instance and a cluster of fine particles are also incorrectly detected as an aggregate (in the small red box). In image (b), all the aggregates are correctly identified. Note also the quality of the segmented masks.

In existing architectures, the classification and detection tasks are done using the same backbone network. In our case, the size of the input images (8092×4096 pixels) is far too large to be used in the neural network with our available GPUs. As mentioned earlier, the classification accuracy depends on the resolution of the images. It is therefore not possible to downscale sufficiently the original image to perform both object localization and classification task. It is not advisable to use patches, because using incomplete image of aggregates would be very detrimental to the mass estimation. We therefore proceed in two steps: a downscaled image of size 1536×768 (resolution of 5.4 pix/mm) is first used as the input of our lightweight instance segmentation network in order to predict the instance bounding boxes and their corresponding masks. Then, each detected aggregate is extracted in the original image (at 28.4 pixels/mm resolution) and resized/padded to a size of 256×256 pixels. Finally, the resized images of individual aggregates are processed by the RACNET. This approach could be improved

509 to make it more efficient (e.g., by integrating the segmentation, classification and mass
510 prediction in the same network), but our aim here is to demonstrate that the proposed
511 approach can be fully automated and used in an industrial environment.

512 **4.2 Validation**

513 To validate our approach, we use a sample of approximately 3kg of RA 6/20 from
514 the recycling platform Valosphere in La Rochelle (France), which belongs to our
515 industrial partner Spie batignolles malet. We first photographed unsorted aggregates
516 using the experimental prototype described in section 2.2, then we manually sorted and
517 weighed the aggregates.

518 The segmentation network extracted 2596 individual aggregates from the 77
519 photographs. Results of manual sorting and RACNET predictions are presented in
520 Table 7: for each class defined by the EN 933-11 standard and as well as for subclasses
521 defined in Table 1. Note that to obtain the classification results, we take the maximum
522 value of the sigmoid of the classification layer output and, if the value is below 0.1, we
523 consider that the aggregate is unknown, and we put it in class X. This threshold value is
524 chosen empirically and is intended to filter predictions with very low confidence only.
525 We also grouped classes X and Rg, as there is only one element in Rg and only a few
526 unknowns.

527 The predictions given in Table 7 are very close to the ground truth for each class
528 defined by the EN 933-11, with a maximum mass fraction deviation of 1.6 point for Rb
529 class (clay bricks, ceramic tiles, earthenware tile, etc.). Furthermore, the relative error
530 between the ground truth and the predicted total mass is only of 2%, which is quite an
531 impressive achievement.

If we look at the detailed subclass predictions in Table 7, we can see that there is however room for improvement. For example, we have found that Ru03 aggregates are not well detected by the network and are often confused with concrete (Rc) or limestone (Ru01) aggregates. We hope that adding more images photographed on the conveyor belt to the D2 database will increase the classification accuracy.

Class	Subclass	M_{GT} (g)	M_{PRED} (g)	%GT	%PRED
Ra		**9.7**	**6.3**	**0.3%**	**0.2%**
Rb		**357.4**	**317.5**	**12.0%**	**10.4%**
	Rb01	299.9	230.5	10.1%	7.6%
	Rb02	57.5	87	1.9%	2.9%
Rc		**1880.1**	**1913.1**	**63.0%**	**62.8%**
Ru		**730.0**	**792.8**	**24.5%**	**26%**
	Ru01	601.0	694.2	20.1%	22.8%
	Ru02	60.3	82	2.0%	2.7%
	Ru03	13.8	2.6	0.5%	0.1%
	Ru04	10.2	1	0.3%	0.03%
	Ru05	44.8	13	1.5%	0.4%
X + Rg		**6.6**	**16.6**	**0.2%**	**0.5%**
TOTAL		**2984**	**3046**	**100%**	**100%**

Table 7: Comparison between manual and automatic classification according to the EN933-11 standard and details for subclasses defined in Table 1.

5 Conclusion and future work

In this work, we have proposed a new method based on convolutional neural networks to estimate the mass and the nature of individual recycled aggregates based on 2D images. We first validated a simple method to estimate the mass of each aggregate on a picture using the total mass of the aggregates and a simple geometric approach,

making possible the creation of a large recycled aggregates database (more than 38k images in database D1 and 20k in database D2).

We designed a new architecture called RACNET to predict the class, the mass and the binary mask of pictures of individual aggregates. The results show that a lightweight network using SepResBlocks allows us to achieve an accuracy of 98% on the test set for the classification task and an average (over classes) error of 3% for the estimation of the total mass of each class. Besides, binary masks areas were predicted with less than 1% error, which will allow further morphological analysis of the detected aggregates (such as particle size distribution).

At last, we demonstrated that our method can be fully automated and applied in realistic conditions, using a simple encoder-decoder network to detect and extract each aggregate. Applied to a real sample coming from a real recycling platform, our method shows very good agreement with manual sorting, even if there are still some intra-class confusions.

The proposed method can therefore allow for rapid characterization and control of recycled aggregates directly in recycling platforms. It is also a first step toward automatic sorting. We hope that this work will help to increase the use of recycled aggregates for the production of higher value-added materials such as concrete.

6 Funding

This work was partly supported by SATT (Tech Transfer Acceleration Companies) Aquitaine Science Transfert.

7 Aknowledgements

We wish to thank Spie batignolles malet for the technical assistance and for providing the necessary materials.

APPENDIX - ResUNet-a architecture and performance

The network used to extract each aggregate is based on the ResUNet-a architecture [35]. The number of filters for each stage is set to 32, 64, 128, 256, 512, 784. The dilation rates in ResUNet -a blocks are set to [1, 3, 9, 15], [1, 3, 9], [1, 3] and [1] depending on the stage. We also add a convolution layer in long range connections between the encoder and decoder part.

The network is trained using the database D2, using 464 images in training set and 50 images in validation set. We use an Adam optimizer with an initial learning rate of 1.10^{-3}, which is halved if loss does not decrease for 3 consecutive epochs. Training is stopped when validation loss does not decrease for 20 consecutive epochs.

The network outputs a contour map, a segmentation map and a distance map (which is not used for instance segmentation). Instances are then extracted as explained in section 4.1.

The network achieves a pixel accuracy of 99,8%, which means that almost all the aggregates pixels are correctly segmented. Regarding the instance detection, the network achieves a box AP_{75} of 96.3% and a box AR_{75} of 97.5% on the validation set.

References

[1] D. Hoornweg, P. Bhada-Tata, What a Waste : A Global Review of Solid Waste Management, World Bank, Washington, DC, 2012. https://openknowledge.worldbank.org/handle/10986/17388.

[2] Transparency Market Research, Construction Waste Market - Global Industry Analysis, Size, Share, Growth, Trends, and Forecast 2017 - 2025, 2021.

https://www.transparencymarketresearch.com/construction-waste-market.html.
[3] Eurostat, Waste statistics, 2021. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Waste_statistics.
[4] ADEME, Déchets chiffres-clés, édition 2020, 2020. https://librairie.ademe.fr/dechets-economie-circulaire/28-dechets-chiffres-cles-edition-2020-9791029712135.html (accessed June 22, 2021).
[5] UNICEM, Granulats : Les Chiffres Clés 2017, 2019. http://www.unicem.fr/wp-content/uploads/stat-unpg-chiffres-2017-web.pdf.
[6] Official Journal of the European Union, Directive 2008/98/EC of the European Parliament and of the Council on waste., 2008. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0098&from=EN.
[7] A. Akbarnezhad, K.C.G. Ong, C.T. Tam, M.H. Zhang, Effects of the Parent Concrete Properties and Crushing Procedure on the Properties of Coarse Recycled Concrete Aggregates, Journal of Materials in Civil Engineering. 25 (2013) 1795–1802. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000789.
[8] M.S. Guimaraes, J.R. Valdes, A.M. Palomino, J.C. Santamarina, Aggregate production: Fines generation during rock crushing, International Journal of Mineral Processing. 81 (2007) 237–247. https://doi.org/10.1016/J.MINPRO.2006.08.004.
[9] S. Omary, E. Ghorbel, G. Wardeh, M.D. Nguyen, Mix Design and Recycled Aggregates Effects on the Concrete's Properties, Int J Civ Eng. 16 (2018) 973–992. https://doi.org/10.1007/s40999-017-0247-y.
[10] M. Bravo, J. de Brito, J. Pontes, L. Evangelista, Mechanical performance of concrete made with aggregates from construction and demolition waste recycling plants, Journal of Cleaner Production. 99 (2015) 59–74. https://doi.org/10.1016/J.JCLEPRO.2015.03.012.
[11] C. Medina, W. Zhu, T. Howind, M. Frías, M.I. Sánchez de Rojas, Effect of the constituents (asphalt, clay materials, floating particles and fines) of construction and demolition waste on the properties of recycled concretes, Construction and Building Materials. 79 (2015) 22–33. https://doi.org/10.1016/J.CONBUILDMAT.2014.12.070.
[12] Y. Hou, Contribution au développement de la valorisation des déchets inertes du BTP : Étude de la carbonatation des granulats recyclés compactés, Thèse de doctorat, 2021. https://www.researchgate.net/publication/354150823_Contribution_au_developpement_de_la_valorisation_des_dechets_inertes_du_BTP_Etude_de_la_carbonatation_des_granulats_recycles_compactes.
[13] N. Zhang, H. Duan, T.R. Miller, V.W.Y. Tam, G. Liu, J. Zuo, Mitigation of carbon dioxide by accelerated sequestration in concrete debris, Renewable and Sustainable Energy Reviews. 117 (2020) 109495. https://doi.org/10.1016/j.rser.2019.109495.
[14] B. Zhan, C.S. Poon, Q. Liu, S. Kou, C. Shi, Experimental study on CO2 curing for enhancement of recycled aggregate properties, Construction and Building Materials. 67 (2014) 3–7. https://doi.org/10.1016/j.conbuildmat.2013.09.008.
[15] J.D. Lau Hiu Hoong, Y. Hou, P. Turcry, P.-Y. Mahieux, H. Hamdoun, O. Amiri, J. Lux, A. Aït-Mokhtar, Reactivity of Recycled Aggregates Used for Pavement Base: From Field to Laboratory, Journal of Materials in Civil Engineering. 33 (2021) 04021129. https://doi.org/10.1061/(ASCE)MT.1943-5533.0003661.
[16] Y. Hou, P.-Y. Mahieux, J. Lux, P. Turcry, A. Aït-Mokhtar, Qualification of the residual reactivity of compacted recycled aggregates, in: ICSBM 2019, 2nd

International Conference of Sustainable Building Materials, Eindhoven, Netherlands, 2019. https://www.researchgate.net/publication/340090865_Qualification_of_the_residual_reactivity_of_compacted_recycled_aggregates.
[17] F. Lancieri, A. Marradi, S. Mannucci, C&D waste for road construction: long time performance of roads constructed using recycled aggregate for unbound pavement layers, in: Waste Management and the Environment III, WIT Press, 2006: pp. 559–569. https://doi.org/10.2495/WM060571.
[18] Y. Hou, X. Ji, L. Zou, S. Liu, X. Su, Performance of cement-stabilised crushed brick aggregates in asphalt pavement base and subbase applications, Road Materials and Pavement Design. 17 (2015) 120–135. https://doi.org/10.1080/14680629.2015.1064466.
[19] A. Grellier, D. Bulteel, M. El Karim Bouarroudj, S. Rémond, Z. Zhao, L. Courard, Alternative hydraulic binder development based on brick fines: Influence of particle size and substitution rate, Journal of Building Engineering. 39 (2021) 102263. https://doi.org/10.1016/j.jobe.2021.102263.
[20] L. Oksri-Nelfia, P.-Y. Mahieux, O. Amiri, Ph. Turcry, J. Lux, Reuse of recycled crushed concrete fines as mineral addition in cementitious materials, Mater Struct. 49 (2016) 3239–3251. https://doi.org/10.1617/s11527-015-0716-1.
[21] J.D. Lau Hiu Hoong, J. Lux, P.-Y. Mahieux, P. Turcry, A. Aït-Mokhtar, Determination of the composition of recycled aggregates using a deep learning-based image analysis, Automation in Construction. 116 (2020) 103204. https://doi.org/10.1016/j.autcon.2020.103204.
[22] L. Zhang, J. Wang, Q. Duan, Estimation for fish mass using image analysis and neural network, Computers and Electronics in Agriculture. 173 (2020) 105439. https://doi.org/10.1016/j.compag.2020.105439.
[23] M. Abdel-Sattar, A.M. Aboukarima, B.M. Alnahdi, Application of artificial neural network and support vector regression in predicting mass of ber fruits (Ziziphus mauritiana Lamk.) based on fruit axial dimensions, PLOS ONE. 16 (2021) e0245228. https://doi.org/10.1371/journal.pone.0245228.
[24] M. Hamdan, D. Rover, M. Darr, J. Just, Mass Estimation from Images using Deep Neural Network and Sparse Ground Truth, in: 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), 2019: pp. 1987–1992. https://doi.org/10.1109/ICMLA.2019.00318.
[25] Y. Miura, Y. Sawamura, Y. Shinomiya, S. Yoshida, Vegetable Mass Estimation based on Monocular Camera using Convolutional Neural Network, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, Toronto, ON, Canada, 2020: pp. 2106–2112. https://doi.org/10.1109/SMC42975.2020.9282930.
[26] T. Standley, O. Sener, D. Chen, S. Savarese, image2mass: Estimating the Mass of an Object from Its Image, in: S. Levine, V. Vanhoucke, K. Goldberg (Eds.), Proceedings of the 1st Annual Conference on Robot Learning, PMLR, 2017: pp. 324–333. https://proceedings.mlr.press/v78/standley17a.html.
[27] K. He, X. Zhang, S. Ren, J. Sun, Identity Mappings in Deep Residual Networks, in: B. Leibe, J. Matas, N. Sebe, M. Welling (Eds.), Computer Vision – ECCV 2016, Springer International Publishing, 2016: pp. 630–645.
[28] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated Residual Transformations for Deep Neural Networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: pp. 5987–5995. https://doi.org/10.1109/cvpr.2017.634.

[29] I. Radosavovic, R.P. Kosaraju, R. Girshick, K. He, P. Dollár, Designing Network Design Spaces, ArXiv:2003.13678 [Cs]. (2020). http://arxiv.org/abs/2003.13678 (accessed December 18, 2020).
[30] M. Tan, Q. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, in: Proceedings of the 36th International Conference on Machine Learning, PMLR, 2019: pp. 6105–6114. https://proceedings.mlr.press/v97/tan19a.html (accessed October 11, 2022).
[31] J. Hu, L. Shen, G. Sun, Squeeze-and-Excitation Networks, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: pp. 7132–7141. https://doi.org/10.1109/CVPR.2018.00745.
[32] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, in: 2017 IEEE International Conference on Computer Vision (ICCV), 2017: pp. 2980–2988. https://doi.org/10.1109/ICCV.2017.322.
[33] Z. Tian, C. Shen, H. Chen, Conditional Convolutions for Instance Segmentation, in: A. Vedaldi, H. Bischof, T. Brox, J.-M. Frahm (Eds.), Computer Vision – ECCV 2020, Springer International Publishing, Cham, 2020: pp. 282–298.
[34] D. Bolya, C. Zhou, F. Xiao, Y. J. Lee, YOLACT++: Better Real-time Instance Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence. (2020) 1–1. https://doi.org/10.1109/TPAMI.2020.3014297.
[35] F.I. Diakogiannis, F. Waldner, P. Caccetta, C. Wu, ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data, ISPRS Journal of Photogrammetry and Remote Sensing. 162 (2020) 94–114. https://doi.org/10.1016/j.isprsjprs.2020.01.013.
[36] F. Waldner, F.I. Diakogiannis, Deep learning on edge: Extracting field boundaries from satellite images with a convolutional neural network, Remote Sensing of Environment. 245 (2020) 111741. https://doi.org/10.1016/j.rse.2020.111741.
[37] G. Yang, Q. Zhang, G. Zhang, EANet: Edge-Aware Network for the Extraction of Buildings from Aerial Images, Remote Sensing. 12 (2020) 2161. https://doi.org/10.3390/rs12132161.

List of Figures

List of Tables