 doi:10.20944/preprints202106.0194.v1

Article

Concrete Cracks Monitoring using Deep Learning-based Multi-resolution Analysis

Ahcene Arbaoui [1], Abdeldjalil Ouahabi [2,3]*, Sébastien Jacques [4] and Madina Hamiane[5]

[1] University of Bouira, Departement of Civil Engineering, Bouira, Algeria; a.arbaoui@univ-bouira.dz (A.A.)
[2] UMR 1253, iBrain, Université de Tours, INSERM, Tours, France;
[3] University of Bouira, Departement of Computer Science, LIMPAF, Bouira, Algeria
[4] University of Tours, GREMAN UMR 7347, CNRS INSA Centre Val-de-Loire, Tours, France; sebastien.jacques@univ-tours.fr (S.J.)
[5] College of Engineering, Royal University for Women, Bahrain mhamiane@ruw.edu.bh

* Correspondence: abdeldjalil.ouahabi@univ-tours.fr (A.O.); Tel.: +33-2-473-613-23

Abstract: In this paper, we propose a new methodology for crack monitoring in concrete structures. This approach is based on a n this paper, we propose a new methodology for monitoring cracks in concrete structures. This approach is based on a multi-resolution analysis of a sample or a specimen of the studied material subjected to several types of solicitation. The image obtained by ultrasonic investigation and processing by a dedicated wavelet will be analyzed according to several scales in order to detect internal cracks and crack initiation. The ultimate goal of this work is to propose an automatic crack type identification scheme based on convolutional neural networks (CNN). In this context, crack propagation can be monitored without access to the concrete surface and the goal is to detect cracks before they are visible on the concrete surface. The key idea allowing such a performance is the combination of two major data analysis tools which are wavelets and Deep Learning. This original procedure allows to reach a high accuracy close to 0.90. In this work, we have also implemented another approach for automatic detection of external cracks by deep learning from publicly available datasets.

Keywords: cracks; wavelets; multiresolution analysis; ultrasound imaging; deep learning; CNN

1. Introduction

To say that concrete is the most widely used man-made material in the world is a no-brainer. Nevertheless, the search for simple, effective and inexpensive techniques to optimize the performance of concrete and to control its mechanical behavior is the real challenge we must meet. In the interests of safety and economy, methods for predicting the performance of concrete structures over the long term (*e.g.* decades) are in great demand, especially in developing countries. Mechanical overload is one of the most frequently cited reasons for damage to concrete. However, other equally devastating factors will be mentioned in this work. Micro-cracks (see, for example, Figure 1 a)) can be caused by excessive mechanical stress, even if this stress is confined to a restricted area. If such stress-overload continues, cracks will continue to form and/or expand, which could lead to excessive damage or even mechanical collapse of the structure. For this reason, crack monitoring is crucial to ensure long-term viability. In current engineering practice, this monitoring is performed by regularly measuring the crack openings at the surface using optical measurements or extensometers. Based on these observations, it is now known that internal damage can lead to leakage or corrosion on large walls, even when only limited cracks are visible on the surface. It is therefore essential to be able to assess and monitor cracks

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

in reinforced concrete constructions at an early stage, especially in special constructions where durability and containment can be significant issues.

Concrete is a mixture of four main materials: Portland cement, coarse aggregate, fine aggregate and water, and for industrial use, mineral and chemicals admixtures are added to accelerate or delay its grip to improve its performance [1, 2]. The quantities of these elements are regulated for a quality required by the destination of the structure's construction, such as long span bridges (see Figure 1 b)), special underground structures (see Figure 1 c)), nuclear power plants (see Figure 2). An excess or defect in the required quantity of one of the constituent elements, inappropriate vibration of the initial mixed elements causes defects such as segregation or premature cracks due to shrinkage of the concrete and the presence of air bubbles causes discontinuities in the material (see Figure 1).

a) b) c)

Figure 1. Examples of cracks in concrete: a) Typical micro-cracking; b) Conventional inspection of bridge cracks; c) Cracks along a slab in a mall.

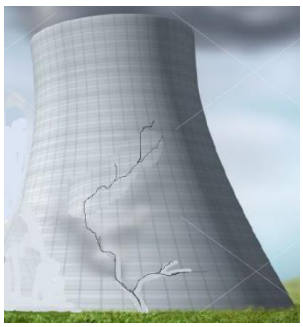

Figure 2. Nuclear power plant: reactor meltdown.

These defects affect the strength of concrete and its durability [3]. Exposed to aggressive environments or temperature variations, visible and non-visible defects appear, its quality and resistance decrease. Under compressive stress, this material behaves well, unlike tensile stresses. In a concrete specimen subjected to compressive stress, the constraints are concentrated on rigid elements with an appreciable modulus of elasticity. Since this material is heterogeneous, an external load creates a complex state within it and a concentration of stresses around air voids [4].

There is a wide variety of methods for evaluating materials or components and non-destructive methods are an important category with multiple applications. The field of non-destructive evaluation (NDE) or non-destructive testing (NDT) [5] involves the identification and characterization of damage to the surface and interior of materials without cutting or other alteration of the material. In other words, NDT refers to the process of evaluating and inspecting materials or components for characterization [6, 7] or searching for defects and flaws in relation to certain standards without altering the original attributes or damaging the test object. NDT techniques provide a cost effective means of testing a sample for individual investigation or can be applied to the entire material for verification in a production quality control system. Thus, NDT is the set of methods that can characterize the state of integrity of structures or materials without degrading them (without

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

altering their function in use). The development of NDT methods began around 1960-1970 to meet the demands of sectors such as nuclear energy, aeronautics and space. NDT gradually widened its field of application, moving from the strict field of detection, recognition and dimensioning of localized defects to the evaluation of the intrinsic characteristics of materials. The notion of defect (or fault) is defined according to the use that will be made of the product (satisfaction of the final customer). NDT methods can be applied to the same elements and structures several times and at different times. These methods are suitable whereby such methods are suitable for diagnostic testing of building structures, both during their construction and during their many years of service.

Detection of cracks is an important task in monitoring the structural health of concrete structures. If cracks develop and continue to propagate, they reduce the effective load bearing surface area and can over time cause failure of the structure.

For this reason, non-destructive testing of concrete now has two main objectives: to detect micro-cracks at an early stage and to monitor stresses in the structures [8, 9].

The main objective of this work is to propose a new approach for the detection of structural cracks in concrete using an ultrasonic non-destructive testing system to scan the concrete and an original methodology based on multi-resolution analysis and deep learning.

The remainder of this paper is organized as follows:

Section 2 is devoted to the foundations of our approach:

- it presents and recalls NDT methods and techniques as well as the experimental set-up used,

- it introduces the main properties of the wavelet transform and the corresponding multiresolution analysis,

- it recalls the foundations of neural networks and CNN-based Deep Learning and proposes the adopted architecture to build a classifier for detecting internal cracks from the obtained spatial-scale images.

Section 5 focuses on the implementation aspect and the analysis of the results.

Finally, Section 6 concludes this study.

2. Materials and Methods

2.1. NDT Methods

The most commonly used NDT methods and techniques [10] are optical NDT [11], ultrasonic NDT [12], acoustic emission testing [13], radiographic NDT [14,15,16], eddy-current testing [17], electromagnetic NDT [18], laser systems [19], sclerometric [20] and thermographic [21] methods for evaluating the durability of concrete structures by contact and non-contact. Figure3 shows the main NDT techniques for concrete structures.

NDT methods are the most desirable and developed methods of concrete diagnosis. A distinction can be made between the stroke method, electrical methods, visual evaluation and acoustic methods. The latter are also called wave methods, based mainly on the analysis of the propagation of ultrasonic waves [22, 23]. Acoustic methods can be divided into passive methods, in which the source of the waves is only a construction with changing load (the acoustic emission method), and active methods of sending and receiving ultrasonic waves. Currently, active methods are not as well developed and tested, so they can be widely used in the field and in most cases require access to two (opposite) sides of the test element or knowledge of its exact dimensions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

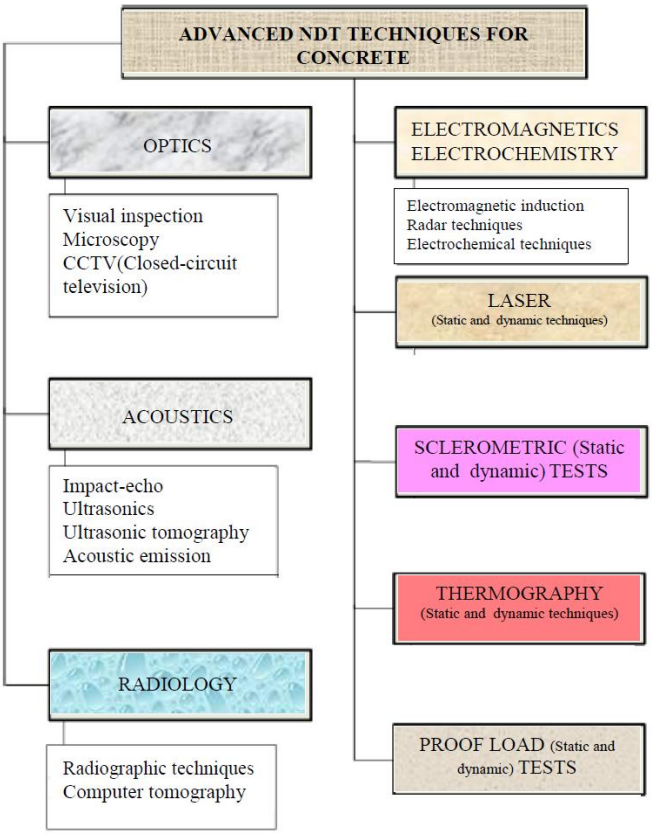

Figure 3. NDT methods and techniques for concrete structures.

Figures 4 and 5 show the devices used experimentally to determine the presence or absence of cracks in concrete subjected to compression. The ultrasound device used is a Pundit L200 from the company Proceq. The press used is a 3R monobloc compression press with a capacity of 2000 kN to 3000 kN adapted to specific tests on concrete specimens of cylindrical shape. The test specimen is of standardized dimension, cylindrical in diameter 16 cm and height 32 cm. Its weight is 15 kg. The test specimen is over 90 days old. The charging speed is 0.05 Mpa / second. The signal transit time varies from 32.3 to 71.4 microseconds when the compressive force varies from 0 to 470 kN.

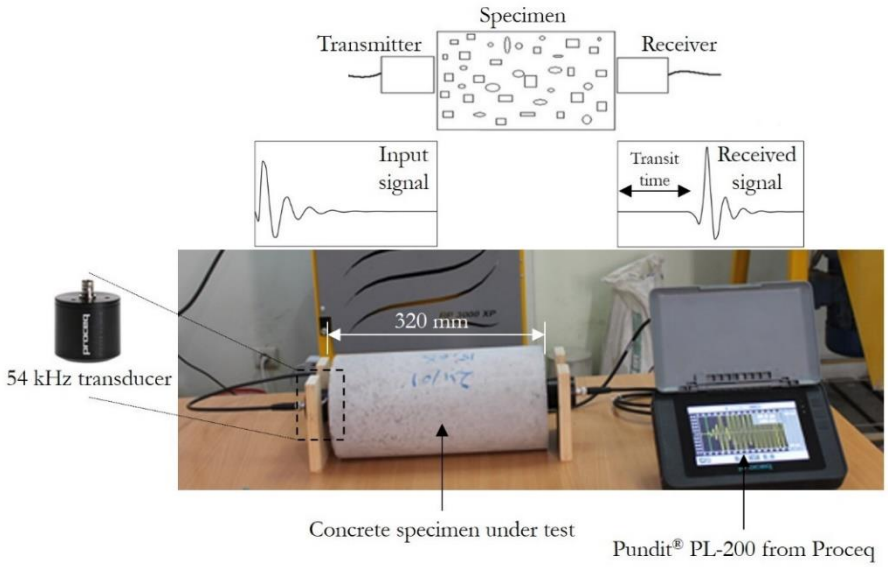

Figure 4. Experimental set-up: ultrasonic non-destructive testing (NDT)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

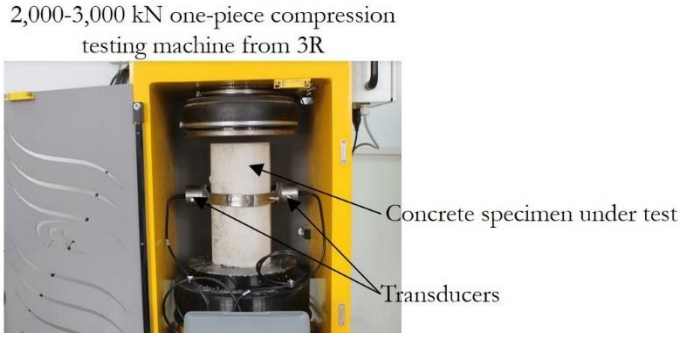

Figure 5. Ultrasonic NDT during compression of a concrete specimen.

The originality of our work lies in the fact that we use, on the one hand, the ultrasound-based NDT method to identify possible cracks and, on the other hand, this method is combined with a multi-resolution analysis based on wavelets to finely analyze cracks and their size at different scales, especially at the beginning of the concrete cracking process. The final objective is to automatically classify these cracks by deep learning and to follow their evolution.

2.2. Multiresolution Analysis Based on Wavelets

The concept of multiresolution analysis [24] provides a framework for the decomposition (and reconstruction) of a signal in the form of a series of approximations of decreasing scale, completed by a series of details. To illustrate this idea, let us take the case of an image constructed from a succession of approximations; the details enhance this image. Thus, coarse vision becomes finer and more precise.

Researchers, engineers and practitioners in various technical fields such as multimedia [25], telecommunications [26-28], medicine and biology [7, 29, 30], crack tracking and fracture detection [8, 9, 31, 32, 33], fluid mechanics [34], thermodynamics [35], astrophysics [36], finance [37, 38], etc., are daily confronted with increasingly challenging technological problems at multiple scales of analysis, in terms of classification, segmentation, detection (of contours or parameters of interest), noise reduction or even elimination [39], compression for transmission or storage, synthesis or reconstruction, etc.

The concept of multiresolution analysis (MRA) is an effective tool that is universally applicable to the above-mentioned fields. This tool, sometimes described as miraculous, produces an immediate and easily interpretable and exploitable result. However, for specific applications that require the extraction of targeted information, it is amply clear that advanced methods will have to be developed and "merged" that exploit existing techniques or optimize analyses (*e.g.* in compression) by taking into account edges or contours, using 2nd and 3rd generation wavelets such as ridgelets [40], curvelets [41], contourlets [42], bandelets [43], etc. Indeed, these anisotropic wavelets are automatically oriented and expanded by unifying the geometry of a potential edge or contour. This conceptualization of multi-resolution analysis is comparable to that of a camera that gets closer to a subject or uses a zoom to distinguish its details, and moves away to capture larger structures - the famous concept of the mathematical microscope.

Figure 4 summarizes the principle of multi-resolution analysis (here for three levels of resolution) based on wavelets. The signal S is first decomposed at the 1st resolution level into an approximation A_1 and a detail D_1, then at the 2nd resolution level, approximation A_1 is decomposed into an approximation A_2 and a detail D_2, and finally at the 3rd resolution level, approximation A_{12} is in turn decomposed into an approximation A_3 and a detail D_3.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

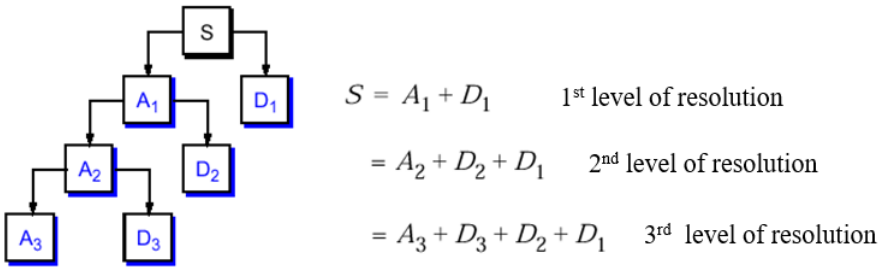

Figure 4. Decomposition of an S signal on three levels. A and D represent respectively the approximation and the detail of the signal.

The signal thus analyzed can be written as follows:

$$S = A_3 + D_3 + D_2 + D_1 \tag{1}$$

Let $\psi(t)$ denote a reference pattern called the mother wavelet. It is generally requested that $\psi(t)$ has jointly highly concentrated time and frequency supports.

$\psi(t)$ satisfies the following equation:

$$\int_{\mathbb{R}} t^p \, \psi(t) dt = 0, \qquad \forall \, 0 \leq p < n \tag{2}$$

where n controls the number of oscillations of $\psi(t)$.

This relation means that $\psi(t)$ is orthogonal to polynomial components of degree less than n.

The wavelet transform $W_X(u,s)$ of a signal X at time u and scale s is defined by (3).

$$W_X(u,s) = \int_{\mathbb{R}} X(t) \, \psi^* \left(\frac{t-u}{s} \right) dt \tag{3}$$

where ψ^*denotes the complex conjugate of ψ.

Looking at expressions (2) and (3), it is clear that $W_X(u,s)$ will be insensitive to the most regular behaviors of the signal assimilated to a polynomial of degree less than n (the number of vanishing moments of ψ). Conversely, $W_X(u,s)$ takes into account the irregular behavior of polynomial tendencies. This important property plays a key role in the detection of signal singularities, especially in the detection and tracking of cracks.

The discrete wavelet transform (DWT) is given by (4).

$$d_X(j,k) = W_X(u = 2^{-j}k, s = 2^{-j}) \quad (j,k) \in \mathbb{Z} \times \mathbb{Z} \tag{4}$$

Clearly, to reduce or eliminate the redundancy, the family $\{\psi_{j,k}\}_{(j,k)\in Z^2}$ must constitute an orthonormal basis of $\mathbb{L}^2(\mathbb{R})$, where $\mathbb{L}^2(\mathbb{R})$denotes the vector space of measurable, square-integrable one-dimensional functions.

This property of the wavelet makes it possible to obtain a fast wavelet transform.

The fast wavelet transformation is calculated by a cascade of low-pass filtering by h and high-pass filtering by g followed by a downsampling (or decimation) by a factor of 2 (see Figure 5).

In Figure 5, a_j (or $a_X(j,k)$, where k represents time) and d_j (or $d_X(j,k)$, where k represents time) are called respectively approximation coefficients and wavelet coefficients (or details) of the signal at level j. Moreover, the symbol $\downarrow 2$ represents the decimation by a factor of 2, in other words, the conservation of one in two samples.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

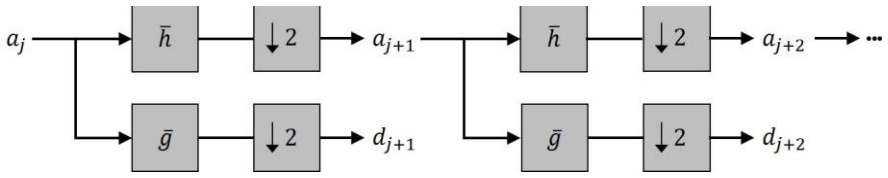

Figure 5. Fast wavelet transform or multiresolution analysis.

The impulse response of the mirror low-pass filter is $\bar{h}(k) = h(-k)$ and that of the mirror high-pass filter is $\bar{g}(k) = g(-k)$. These two impulse responses [10] are linked by $g(k) = (-1)^k h(1-k)$ whose coefficients are obtained directly from the chosen wavelet ψ.

In Figure 6, it should be noted that the original signal has 1,000 samples while the detail (and approximation) signals have been decimated by a factor of 2 at each level of resolution. Hence, after 3 levels of resolution, from a signal of 1,000 samples, one arrives at the approximation A_3 and the detail D_3 which each have only 125 samples.

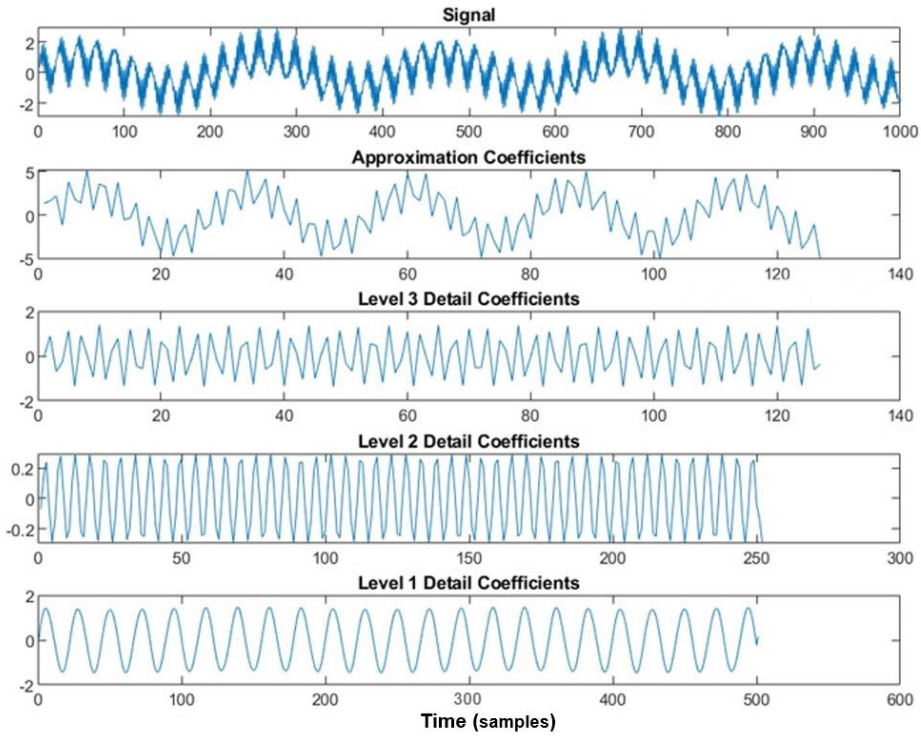

Figure 6. Signal decomposition on three levels of resolution using MATLAB.

In this study, the scalogram of the investigative ultrasound signal will be used to determine and analyze cracks in concrete. We can define the scalogram of the signal $x(t)$ by

$$S_X(j,k) = |d_X(j,k)|^2 \quad \forall\, (j,k) \in \mathbb{Z} \times \mathbb{Z} \tag{5}$$

Figure 7 shows the scalogram of a signal representing the initialization of a crack materialized by intense energy. This fracture also propagates, even if in a weaker way, to other scales which can cause in the long term to a rupture.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

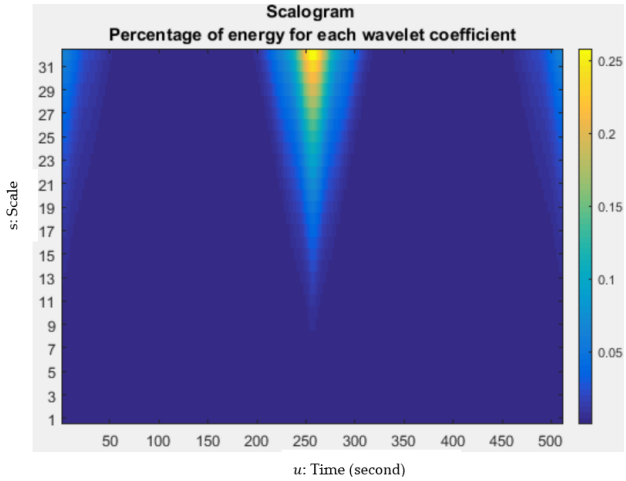

Figure 7. Scalogram of a signal representing the initialization of a crack

2.3. From neurons to CNN and Deep Learning: Basic concepts

2.3.1. Neuronal Networks

Neural networks is one of the most popular machine learning algorithms at present. It has been decisively proven over time that neural networks outperform other algorithms in accuracy and speed. With various variants like CNN (convolutional neural networks), neural networks are becoming for data scientists or machine learning practitioners what linear regression was one for statisticians. It is thus imperative to have a fundamental understanding of what a neural network is, how it is made up and what is its reach and limitations.

Neural networks are modeled as collections of neurons that are connected in an acyclic graph. In other words, the outputs of some neurons can become inputs to other neurons. For classic neural networks, the most common type of layer is the fully connected layer where all inputs from one layer are connected to each activation unit of the next layer. In most common machine learning models, the last layers are fully connected layers that compile the data extracted by the previous layers to form the final output. In Fig. 5, two example neural network topologies that use a stack of fully-connected layer.

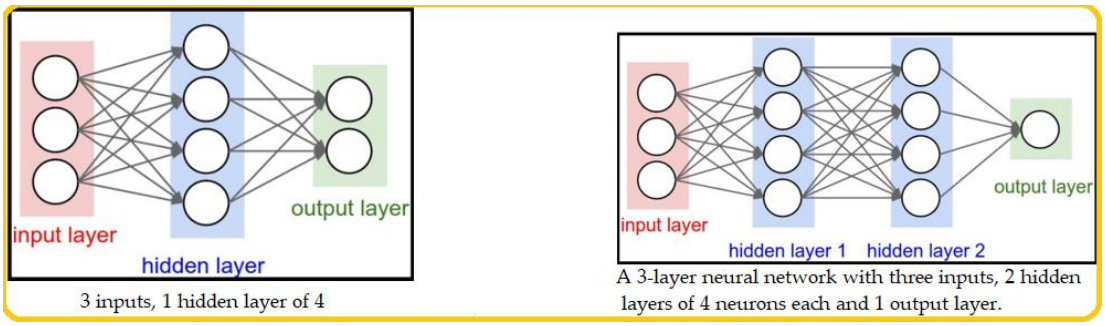

Figure 8. Neural networks

The two metrics that people commonly use to measure the size of neural networks are the number of neurons, or more commonly the number of parameters. Working with the two example networks in Figure 8:

The first network (left) has $4 + 2 = 6$ neurons (not counting the inputs), $[3 \times 4 + 4 \times 2] = 20$ weights and $4 + 2 = 6$ biases, for a total of 26 learnable parameters.

The second network (right) has $4 + 4 + 1 = 9$ neurons, $3 \times 4 + 4 \times 4 + 4 \times 1 = 12 + 16 + 4 = 32$ weights and $4 + 4 + 1 = 9$ biases, for a total of 41 learnable parameters.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

Traditional neural networks use a fully-connected architecture, as illustrated in Figure 8 (left), where every neuron in one layer connects to all the neurons in the next layer. A fully connected architecture is inefficient when it comes to processing image data as in our case. For an average image with hundreds of pixels and three channels red, green and blue (RGB), a traditional neural network will generate millions of parameters, which can lead to overfitting. The model would be very computationally intensive and it may be difficult to interpret results, debug and tune the model to improve its performance.

Modern convolutional networks contain on orders of 100 million parameters and are typically made up of around 10 to 20 layers (hence deep learning). However, as we will see, the number of effective connections is significantly higher due to parameter sharing.

2.3.2. CNN and Deep Learning

Deep learning is the new state-of-the-art for artificial intelligence. Deep learning architecture is composed of an input layer, hidden layers, and an output layer. The word deep means there are more than two fully connected layers. Convolutional networks are a specialized type of neural networks that use convolution in place of general matrix multiplication in at least one of their layers.

Convolutional Neural Network (CNN) is one of the main categories to do images recognition, images classifications [44-47], crack damage detection [48]. The name Convolutional Neural Network indicates that the network uses a mathematical operation called convolution which performs a filtering of the original image by a filter or a kernel in order to extract features.

There are six main operations in the CNN architecture shown in Figure 9:

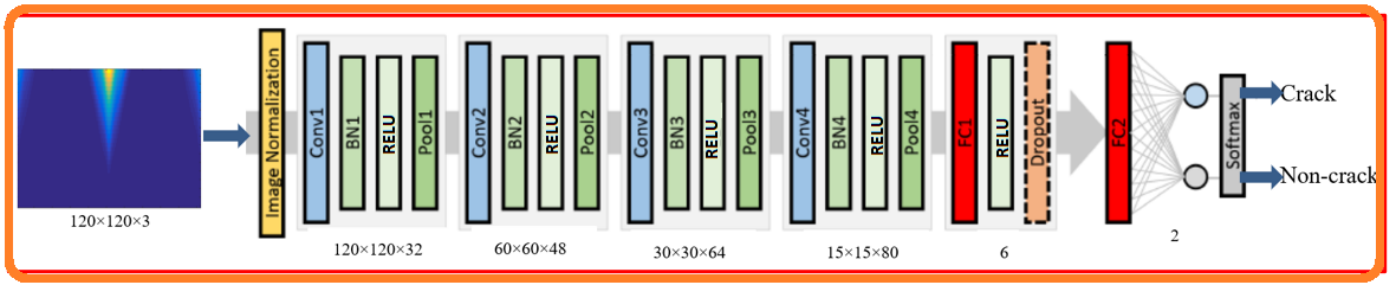

Figure 9. CNN architecture (four layers) for crack and non-crack concrete classification. The numbers below layers indicate the output data size of each convolution or fully-connected layer.

1. **CONV**: Convolution. The convolution layer is the core building block of the CNN. It carries the main portion of the network's computational load. This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.

2. **BN**: Batch normalization is a technique to standardize the inputs to a network, applied to ether the activations of a prior layer or inputs directly. Batch normalization accelerates training, in some cases by halving the epochs or better, and provides some regularization, reducing generalization error.

3. **RELU**: Rectified Linear Unit. Since convolution is a linear operation and images are far from linear, non-linearity layers are often placed directly after the convolutional layer to introduce non-linearity to the activation map. There are several types of non-linear operations, the ReLU has become very popular in the last few years. It computes the function $y = \max(0, x)$ (see Figure 10). In other words, the activation is simply threshold at zero.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

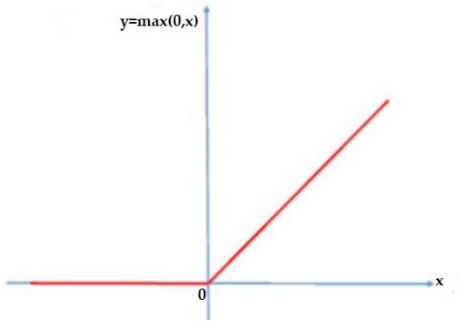

Figure 10. Rectified Linear Unit (RELU)

4. **Pool**: Pooling or Sub Sampling. The pooling layer replaces the output of the network at certain locations by deriving a summary statistic of the nearby outputs. This helps in reducing the spatial size of the representation, which decreases the required amount of computation and weights. The pooling operation is processed on every slice of the representation individually. There are several pooling functions such as the average of the rectangular neighborhood, L2 norm of the rectangular neighborhood, and a weighted average based on the distance from the central pixel. However, the most popular process is max pooling, which reports the maximum output from the neighborhood.
5. **FC**: Fully Connected Layer. Neurons in this layer have full connectivity with all neurons in the preceding and succeeding layer as seen in regular FCNN. This is why it can be computed as usual by a matrix multiplication followed by a bias effect. The FC layer helps to map the representation between the input and the output.
6. **Dropout** is a regularization technique for neural network models where randomly selected neurons are ignored during training. This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass. As a neural network learns, neuron weights settle into their context within the network. Weights of neurons are tuned for specific features providing some specialization. Neighboring neurons become to rely on this specialization, which if taken too far can result in a fragile model too specialized to the training data. This reliant on context for a neuron during training is referred to complex co-adaptations. You can imagine that if neurons are randomly dropped out of the network during training, that other neurons will have to step in and handle the representation required to make predictions for the missing neurons. This is believed to result in multiple independent internal representations being learned by the network. The effect is that the network becomes less sensitive to the specific weights of neurons. This in turn results in a network that is capable of better generalization and is less likely to overfit the training data.

7. **Softmax** is implemented via a neural network layer just before the result layer. The Softmax layer must have the same number of nodes as the result layer. In probability theory, the output of the softmax function can be used to represent a categorical distribution - that is, a probability distribution over K different possible outcomes representing the possible classes in classification. In our application, we have 2 possible cases: For example $p_1 = 0.95$ for crack and $p_2 = 0.05$ for non-crack.

Figure 9 shows a CNN architecture adapted to our problem of monitoring concrete structures. This consists of 4 CONV layers, 4 BN layers, 4 ELU layers and 4 Pool layers, followed by a FC, RELU and Dropout layers. Finally, an FC layer decides, via Softmax activation, the final classification of the image into crack or non-crack.

3. Results and Discussion

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

3.1. Metrics and Data

In this paragraph, we will define the metrics used to evaluate the performance of our Deep Learning approach, and explain that there are two methodologies to adopt depending on the nature of the concrete crack: internal or external.

To achieve our goal of classifying cracked/non-cracked concrete images, it is necessary to use evaluation measures to assess the performance of our approach. Accuracy is the ratio of the number of correctly predicted cracked and uncracked images to the total number of input images. Accuracy is the most intuitive one and is defined as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

where where TP (True Positive) and TN (True Negative) mean images with crack and without crack, which are correctly classified. FP (False Positive) and FN (False Negative) mean images with crack and without crack which are wrongly classified.

Precision aka confidence or true positive accuracy (positive prediction value) can be understood as the number of correctly predicted crack images divided by the number of crack images predicted by the classifier. Precision can be interpreted as an indicator of robustness. It is defined as:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall aka sensitivity or true positive rate is the percentage of the number of correctly predicted crack images to the total number of crack images.

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

F_β score is a weighted harmonic average comprehensively reflecting the $Precision$ and $Recall$. It is defined as:

$$F_\beta = (1 + \beta^2)\frac{Precision \times Recall}{\beta^2 Precision + Recall} \tag{9}$$

where β is a coefficient to trade off precision and recall. β is set to be 1 here to give the precision rate and recall rate the same weight. In this context, F_1 score is the harmonic mean of $Precision$ and $Recall$ and it is defined as:

$$F_1 = 2\frac{Precision \times Recall}{Precision + Recall} = 2\frac{TP}{2TP + FP + FN} \tag{10}$$

In this work, we will use two sources of images of concrete cracks:

1. The first source is derived from ultrasonic non-destructive testing images of internal cracks analyzed by wavelets. The multiresolution images are then classified into cracks/no cracks by deep learning.

The procedure for these images is described in Section 2. This is our main contribution here.

2. The second source of images comes from SDNET2018 dataset.

SDNET2018 is anannotated image dataset fort raining,validation, and benchmarking of artificial intelligence based crack detection algorithms for concrete. SDNET2018 contains over 56,000 images of cracked and non-cracked concrete bridge decks, walls, and pavements (see Figure 11). The dataset includes cracks as narrow as 0.06 mm and as wide as 25mm. The dataset also includes images with a variety of obstructions, including shadows, surface roughness, scaling, edges, holes, and background debris (see Figure 12).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

SDNET2018 will be useful for the continued development of concrete crack detection algorithms based on deep convolutional neural networks.

Figure 11. Noncrack and crack images form SDNET2018 dataset

Figure 12. SDNET2018 images include (a) fine cracks, (b) coarse cracks, (c) shadows, (d) stains, (e) rough surface finishes, (f) inclusion sand voids, (g) edges, (h) joint sand surface scaling, and (i) background obstructions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

3.2. *Implementation aspect and results analysis*

For both data sources, i.e., experimentally obtained NDT and wavelet-transformed images, and images from SDNET2018 dataset [49], we selected 1000 images with cracks and 1000 images without cracks. The size of each image is 256×256 pixels RGB.

For the SDNET2018 dataset, these are images of concrete bridges and it was possible to introduce various changes, such as changes in lighting conditions and crack characteristics as well as crack surface texture in order to further test the generalizability of the model and make a more comprehensive evaluation of the model.

All experiments are run on TensorBook equipped with a Super GPU 2080. It is equipped with Lambda Stack which is a software tool to manage installations of TensorFlow, Keras, PyTorch, Caffe, Caffe 2, Theano, CUDA and cuDNN. Its specifications are :

Display: 16.1″ FHD (1920×1080) display, matte finish.

Storage: 1TB NVME SSD

RAM: 64 GB (2666 MHz)

Processor: Intel i7-8750H (6 cores, 16x PCI-e lanes)

Graphics: NVIDIA 2080 (8GB)

Operating System: Ubuntu 18.04 (Bionic) with Lambda Stack and Windows 10 Pro also available

In this high performance laptop all TensorFlow Keras Pytorch Caffe Theano etc. frameworks specially designed for deep learning are pre-installed. It is optimized for performing TensorFlow training tasks with its powerful Pascal GTX.

In order to have a comparative evaluation of the performance of the deep learning architecture to adopt, we used both the version of AlexNet [50] represented in figure 9 and an architecture of the ResNet type [51] . ResNet, short for Residual Networks, is a classical neural network used as the backbone for many computer vision tasks. This model was the winner of the ImageNet competition in 2015. The fundamental advance of ResNet is that it successfully trains extremely deep neural networks with over 150 layers. Before ResNet, training extremely deep neural networks was difficult due to the problem of evanescent gradients. AlexNet, the winner of ImageNet 2012 and the model that apparently kicked off deep learning, had only 8 convolutional layers, the VGG network had 19, Inception or GoogleNet had 22 and ResNet 152 had 152. In our work, we will code a ResNet-50 which is a reduced version of ResNet 152 and is frequently used as a starting point for transfer learning.

ResNet is a powerful backbone model that is used very frequently in many computer vision tasks. Its originality is to use skip connection to add the output from an earlier layer to a later layer. This helps it mitigate the vanishing gradient problem.

In this study, we used Keras to load their pre-trained ResNet 50 (see Figure 13).

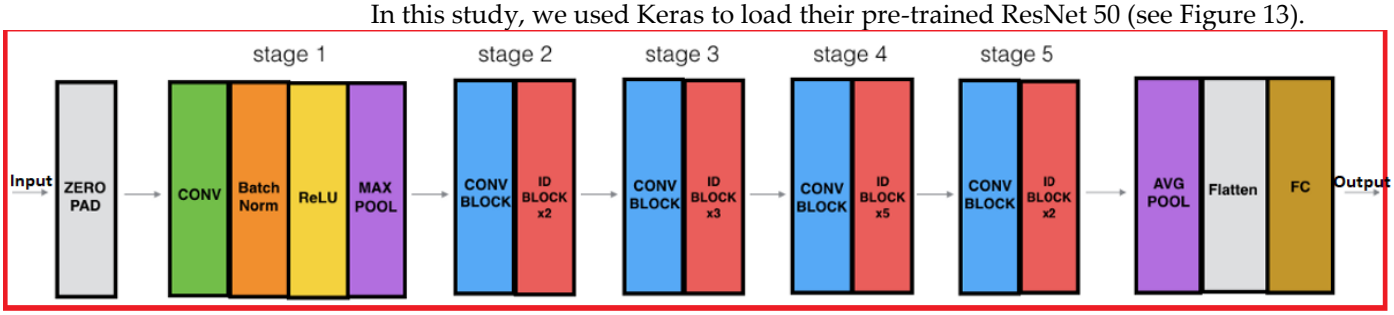

Figure 13. ResNet 50 Model

ResNet-50 model consists of 5 stages each with a convolution and identity block. Each convolution block has 3 convolution layers and each identity block also has 3 convolution layers. The ResNet-50 has over 23 million trainable parameters.

Classical Adam optimization of stochastic gradient descent is used for training [52].

Table 1 shows the results from the NDT procedure based on multiresolution analysis to detect internal cracks in a concrete structure.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 June 2021 doi:10.20944/preprints202106.0194.v1

Table 2 shows the results from SDNET2018 dataset.

In both tables, two deep learning architectures are compared in terms of Accuracy, Precision, Recall and F1 score.

Table 1. NDT-Multiresolution Analysis: Comparison of two model's crack detection results.

Model	Accuracy	Precision	Recall	F_1
AlexNet	0.8855	0.8921	0.8840	0.8881
ResNet50	0.9068	0.9178	0.9091	0.9099

Table 2. SDNET2018: Comparison of two model's crack detection results.

Model	Accuracy	Precision	Recall	F_1
AlexNet	0.9182	0.9322	0.9241	0.9284
ResNet50	0.9691	0.9691	0.9704	0.9798

Table 1 shows that the performance of the ResNet50 architecture is superior to that of AlexNet. This was to be expected but the difference is not very pronounced. It should be noted that the Accuracy from the method proposed here which is the detection of internal cracks from NDT followed by a wavelet-based multiresolution analysis is capped at 90%.

On the other hand, Table 2 shows high performance and the difference between the ResNet50 and AlexNet architecture is more clear.

The apparent limitation of the NDT-Mutiresolution Analysis method is explained by the fact that the crack is internal and is more difficult to detect than a surface crack. In reality, the method we propose is very efficient since it allows the detection of an invisible crack by optical means which would avoid many disasters in sensitive structures.

4. Conclusions

In this work, we proposed an original method for monitoring cracks in concrete structures. This method focuses on internal cracks or on the beginning of cracks invisible from the outside.

Such cracks are detected by ultrasonic NDT and analyzed by wavelets providing a space-scale image allowing to localize the crack in space and at each resolution.

The resulting multiresolution image is then subjected to a crack/non-crack classification process based on Deep Learning (AlexNet, ResNet).

We have shown that it is possible to reach an accuracy of 90%.

Such a result is very positive and shows that our approach is unavoidable when it comes to "securing" vital economic structures such as nuclear power plants and dams where the initialization of an optically invisible crack can cause major disasters.

Author Contributions: Data curation, conceptualization, resources, project administration and funding acquisition A. A.; methodology, investigation, validation and writing—review and editing A.0.; software, formal analysis, S.J.; visualization, supervision, M.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hassoum, M.N.H.; El-Manaseer, A. *Structural concrete: Theory and design*, John Wiley & Sons, 2015.
2. Bernier, G., *Formulation des bétons*, Techniques de l'ingénieur. Réf. : C2210 V2, 2004.
3. Mehta, P. K.; Monteiro P. J. M. *Concrete, Microstructure, Properties and Materials*, The McGraw-Hill Companies, 2006.
4. Mehta, P. K. Durability – Critical Issues for the future. *Concrete International* **1997**, *9*, 27-33.

5. Verma, S. K.; Bhadauria, S. S.; Akhtar, S. Review of Nondestructive testing methods for condition monitoring of Concrete Structures. *Journal of Construction Engineering* **2013**, *2013*, 11.
6. Femmam, S. ; M'sirdi N. K. ; Ouahabi A. Perception and characterization of materials using signal processing techniques. *IEEE Transactions on Instrumentation & Measurements* **2001**, *50*, *5*, 1203-1211.
7. Ouahabi, A. Multifractal analysis for texture characterization: A new approach based on DWT. In Proceedings of the 10th International Conference on Information Science, Signal Processing and their Applications (IEEE/ISSPA), Kuala Lumpur, Malaysia, **2010**, 698-703.
8. Ait Aouit, D. ; Ouahabi A. Monitoring crack growth using thermography. *Comptes Rendus Mécanique* **2008**, *336*, *8*, 677-683.
9. Ait Aouit, D.; Ouahabi, A. Nonlinear Fracture Signal Analysis Using Multifractal Approach Combined with Wavelet. *Fractals - Complex Geometry, Patterns, and Scaling in Nature and Society* **2011**, *19*, *18*, 175-183.
10. Denys, B. Nondestructive evaluation of concrete strength: An historical review and a new perspective by combining NDT methods. *Construction and Building Materials* **2012**. *33*, 139–163.
11. Zhu, Y.-K.; Tian, G.-Y.; Lu, R.-S.; Zhang, H. A Review of Optical NDT Technologies. *Sensors* **2011**, *11*, 7773-7798.
12. Girault, J.-M.; Ossant, F.; Ouahabi, A.; Kouame, D.; Patat, F. Time-varying autoregressive spectral estimation for ultrasound attenuation in tissue characterization. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* **1998**, *45*, *3*, 650-659.
13. Zhao, L.; Kang, L.; S. Yao, S. Research and Application of Acoustic Emission Signal Processing Technology. *IEEE Access* **2019**, *7*, 984-993.
14. Liao, T. W.; Ni, J. An automated radiographic NDT system for weld inspection: Part I — Weld extraction. NDT & E International **1996**, *29*, *3* ;157-162.
15. Tang, S.; Chris Ramseyer, C.; Pratik Samant, P.; Liangzhong Xiang, L. X-ray-induced acoustic computed tomography of concrete infrastructure. *Applied Physics Letters* **2018**, *112*, 063504.
16. Zehani, S.; Ouahabi, A.; Oussalah, M.; Mimi, M.; Taleb-Ahmed, A. Trabecular bone microarchitecture characterization based on fractal model in spatial frequency domain imaging. *Int. J. Imaging Syst. Technol.* **2021**, *31*, 141-159.
17. Chady, T. ; Enokizono, M. ; Sikora, R. Crack detection and recognition using an eddy current differential probe. *IEEE Transactions on Magnetics* **1999**, *35*, *3*, 1849-1852.
18. Langenberg, K.-J; Mayer, K.; Marklein, R. Nondestructive testing of concrete with electromagnetic and elastic waves: Modeling and imaging. *Cement and Concrete Composites* **2006**, *28*, *4*, 370-383.
19. Ouahabi, A.; Depollier, C.; Simon, L.; Kouame, D. Spectrum estimation from randomly sampled velocity data [LDV]. *IEEE Trans. Instrum. Meas* **1998**, *47*, *4*, 1005–1012.
20. Kowalski, R. ; Wróblewska, J. Application of a sclerometer to the preliminary assessment of concrete quality in structures after fire. Archives of Civil Engineering **2018**, *64*, *4*, 171-186.
21. Foudazi, A. ; Edwards, C. A. ; Ghasr, M. T. ; Donnell, K. M. Active Microwave Thermography for Defect Detection of CFRP-Strengthened Cement-Based Materials. *IEEE Transactions on Instrumentation and Measurement* **2016**, *65*, *11*, 2612-2620.
22. Girault, J.-M.; Kouamé, D.; Ouahabi, A.; Patat, F. Estimation of the blood Doppler frequency shift by a time-varying parametric approach. *Ultrasonics* **2000**, *38*, *1*, 682-687.
23. Labat, V.; Remenieras, J.-P.; Matar, O.B.; Ouahabi, A.; Patat, F. Harmonic propagation of finite amplitude sound beams: experimental determination of the nonlinearity parameter B/A. *Ultrasonics* **2000**, *38*, *1-8*, 292-296.
24. Ouahabi, A. *Signal and Image Multiresolution Analysis*, London (UK)/Hoboken (NJ-US), ISTE-Wiley, 2013.
25. Sui, K.; Kim, HG. Research on application of multimedia image processing technology based on wavelet transform. *Image Video Processing* **2019**, *24*.
26. Haneche, H.; Ouahabi, A.; Boudraa, B. New mobile communication system design for Rayleigh environments based on compressed sensing-source coding. *IET Communications* **2019**, *13*, *15*, 2375–2385.
27. Haneche, H.; Boudraa, B.; Ouahabi, A. A new way to enhance speech signal based on compressed sensing, *Measurement* **2020**, *151*, 107–117.
28. Haneche, H.; Ouahabi, A.; Boudraa, B. Compressed Sensing-Speech Coding Scheme for Mobile Communications. *Circuits, Systems, and Signal Processing* **2021**.
29. Djeddi, M.; Ouahabi, A.; Batatia, H.; Basarab, A.; Kouamé, D. Discrete wavelet transform for multifractal texture classification: Application to ultrasound imaging. In Proc. of the 17th IEEE International Conference on Image Processing (ICIP), 2010, Hong Kong, 637-640.
30. Guetbi, C.; Kouamé, D.; Ouahabi, A.; Remenieras, J.-P. New emboli detection methods [Doppler ultrasound]. In 1997 IEEE Ultrasonics Symposium Proceedings, An International Symposium (Cat. No.97CH36118), Toronto, Ont., Canada, 1997; pp. 1119-1122.
31. Calvez, D. ; Roqueta, F. ; Jacques, S. ; Béchou, L.; Ousten, Y. ; Ducret, S. Crack Propagation Modeling in Silicon: a Comprehensive Thermo-Mechanical FEM Approach for Power Devices. *IEEE Transactions on Components Packaging Manufacturing Technologies* **2014**, *4*, *2*, 360-366.
32. Ouahabi, A.; Femmam, S. Wavelet-based multifractal analysis of 1-D and 2-D signals: new results. *Analog Integrated Circuits and Signal processing* **2011**, *69*, *1*, 3-15.
33. Girault, J.-M.; Kouamé, D.; Ouahabi, A. Analytical formulation of the fractal dimension of filtered stochastic signal. *Signal Processing* **2010**, *90*, *9*, 2690-2697.
34. Schneider K.; Vasilyev O. V. Wavelet methods in computational fluid dynamics. Annual Review of Fluid Mechanics. *Annual Reviews* **2010**, *42*, 473-503.

35. Arneodo, A.; Bacry, E.; Muzy, J. F. The thermodynamics of fractals revisited with wavelets. *Physica A: Statistical Mechanics and its Applications* **1995**, *213*, *1-2*, 232-275.
36. De Moortel, I.; Hood, A.W. Wavelet analysis and the determination of coronal plasma properties. *Astronomy and astrophysics* **2000**, *363*, 269–278.
37. Ramsey, J. B. Wavelets in Economics and Finance: Past and Future. *Studies in Nonlinear Dynamics and Econometrics* **2002**. *6*, *3*.
38. Struzik, Z. R. Wavelet methods in (financial) time-series processing, *Physica A: Statistical Mechanics and its Applications* **2001**, *Volume 296*, *Issues 1–2*; pp. 307-319.
39. Ouahabi, A. A Review of Wavelet Denoising in Medical Imaging. In Proceedings of the 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA), Tipaza, Algeria, 12–15 May 2013, 19–26.
40. Do, M. N.; Vetterli, M. The finite ridgelet transform for image representation. *IEEE Transactions on Image Processing* **2003**, *12*, *1*, 16-28.
41. Fadili, J.M.; Starck, J.-L. Curvelets and ridgelets, *Encyclopedia of Complexity and Systems Science* **2009**, *Volume 3*; pp. 1718–1738.
42. Sidahmed, S.; Messali, Z.; Ouahabi, A.; Trépout, S.; Messaoudi, C.; Marco, S. Nonparametric Denoising Methods Based on Contourlet Transform with Sharp Frequency Localization: Application to Low Exposure Time Electron Microscopy Images. *Entropy* **2015**, *17*, 3461–3478.
43. Ferroukhi, M.; Ouahabi, A.; Attari, M.; Habchi, Y.; Taleb-Ahmed, A. Medical video coding based on 2nd-generation wavelets: Performance evaluation. *Electronics* **2019**, *8*, 88.
44. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Taleb-Ahmed, A. Past, Present, and Future of Face Recognition: A Review. *Electronics* **2020**, *9*, 1188.
45. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Jacques, S. Multi-Block Color-Binarized Statistical Images for Single-Sample Face Recognition. *Sensors* **2021**, *21*, 728.
46. Ouahabi, A.; Taleb-Ahmed. A. Deep learning for real-time semantic segmentation: Application in ultrasound imaging. *Pattern Recognition Letters* **2021**, *144*, 27–34.
47. Khaldi, Y.; Benzaoui, A.; Ouahabi, A.; Jacques, S.; Taleb-Ahmed, A. Ear Recognition Based on Deep Unsupervised Active Learning. IEEE Sensors Journal **2021**, Accepted.
48. Cha, Y.; Choi, W.; Büyüköztürk, O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering* **2017**, *32*; 361-378.
49. Dorafshan, S. ; Thomas, R.J. ; Maguire. M. SDNET2018: an annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks. *Data in Brief* **2018**, *21*, 1664–1668.
50. Krizhevsky, A.; Sutskever, I.; Hinton, G. *ImageNet classification with deep convolutional neural networks". Communications of the ACM* **2017**, *60*, *6*, 84–90.
51. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; 770–778.
52. Kingma, D. P.; Ba, J. Adam: a method for stochastic optimization. *arXiv* **2014**, 1412. 6980.