Article

Depth Data Denoising in Optical Laser Based Sensors for Metal Sheet Flatness Measurement: A Deep Learning Approach

Marcos Alonso [1,2,*], **Daniel Maestro** [1], **Alberto Izaguirre** [1], **Imanol Andonegui** [1] **and Manuel Graña** [2]

1   Robotics and Automation Group, Electronic and Computer Science Department, Faculty of Engineering, Mondragon University, Loramendi Kalea, 4, 20500 Arrasate-Mondragon, Spain; dmaestro@mondragon.edu (D.M.); aizagirre@mondragon.edu (A.I.); iandonegui@mondragon.edu (I.A.)
2   Computational Intelligence Group, CCIA Department, UPV/EHU, Paseo Manuel de Lardizabal 1, 20018 San Sebastian, Spain; manuel.grana@ehu.es
*   Correspondence: malonson@mondragon.edu

Abstract: Surface flatness assessment is necessary for quality control of metal sheets manufactured from steel coils by roll leveling and cutting. Mechanical-contact-based flatness sensors are being replaced by modern laser-based optical sensors that deliver accurate and dense reconstruction of metal sheet surfaces for flatness index computation. However, the surface range images captured by these optical sensors are corrupted by very specific kinds of noise due to vibrations caused by mechanical processes like degreasing, cleaning, polishing, shearing, and transporting roll systems. Therefore, high-quality flatness optical measurement systems strongly depend on the quality of image denoising methods applied to extract the true surface height image. This paper presents a deep learning architecture for removing these specific kinds of noise from the range images obtained by a laser based range sensor installed in a rolling and shearing line, in order to allow accurate flatness measurements from the clean range images. The proposed convolutional blind residual denoising network (CBRDNet) is composed of a noise estimation module and a noise removal module implemented by specific adaptation of semantic convolutional neural networks. The CBRDNet is validated on both synthetic and real noisy range image data that exhibit the most critical kinds of noise that arise throughout the metal sheet production process. Real data were obtained from a single laser line triangulation flatness sensor installed in a roll leveling and cut to length line. Computational experiments over both synthetic and real datasets clearly demonstrate that CBRDNet achieves superior performance in comparison to traditional 1D and 2D filtering methods, and state-of-the-art CNN-based denoising techniques. The experimental validation results show a reduction in error than can be up to 15% relative to solutions based on traditional 1D and 2D filtering methods and between 10% and 3% relative to the other deep learning denoising architectures recently reported in the literature.

Keywords: laser triangulation; metal sheet flatness measurement; smooth surface reconstruction; depth data denoising; Convolutional Neural Networks; deep learning, residual learning

1. Introduction

Increasingly stringent specifications in terms of flatness and surface quality in the manufacture of sheet metal products are becoming more demanding of real-time 100% quality-control processes. The end customer expects not only excellent mechanical and processing properties but also a high long-term value and a high quality of modern metallic materials. To meet these high expectations, the steel industry needs intelligent quality-control systems endowed with high-precision in-line sensors for real-time measurements.

In the manufacture of parts and assemblies, especially when parts are required to be assembled over a surface, flatness is a critical specification requirement. Any flatness defect will cause an undesirable optical effect and impact the overall appearance of the assembly. This need for zero defect manufacturing arises in areas as varied as the manufacture of

stainless steel sheets used in professional kitchens, metal panels for exterior decoration in architectural projects, or the manufacture of aluminum sheets in the automotive industry. For this reason, it is highly desirable to carry out a quality control in real time during the metal sheet manufacturing covering 100% of its surface in order to ensure that the required industry quality standards are met.

With the advances in computer vision technology, optical flatness sensors have became widespread [1] allowing manufacturing line human operators to measure manifest flatness, i.e., flatness not hidden by tension, at high line speeds, thus enabling real-time monitoring as well as a high degree of automation in the production phase. Most optical surface flatness inspection systems used in the metal sheet industry are based on the laser triangulation principle [2–4].

The large real-time inspection capabilities of these optical sensors are impeded by the non-linear high-frequency fluctuations induced in the steel sheet surface by the mechanical processes that take place in the manufacturing line, the juddering of the metal strip due to forward traction, as well as the shearing processes that cut to length the sheet. Under these circumstances, achieving a highly accurate flatness measurement requires a high performance signal denoising method to be applied to the height profile captured by the 3D sensor, removing the noise corresponding to such non-linear high-frequency fluctuations. The literature [5–11] presents different sensors based on laser triangulation, requiring the use of two or more laser lines to filter out external noise sources and reconstruct an accurate and smooth continuous 3D map of the metal sheet surface.

The main contribution of this paper is a novel deep learning architecture for the reconstruction of the range image captured by the 3D sensor removing the high-frequency noise due to mechanical processes in order to allow accurate flatness measurements for quality control. This deep learning architecture is inspired in the U-Net [12], originally developed for semantic segmentation. Instead of returning as output an decomposition of the image into regions, our architecture returns the noise-free range image by using a noise estimation module. The architecture is validated against synthetic and real range images that exhibit the most significant noise modalities produced by the mechanical processing induced vibrations on the steel sheet surface. Real data have been collected from an industrial roll leveling and cut-to-length line where the developed 3D sensor is installed. Moreover, the architecture is compared against image denoising deep learning architectures reported in the literature. To this end, we have retrained these architectures with our data from scratch.

The remainder of this paper is organized as follows: Section 2 reviews the industrial context regarding techniques and devices used to measure metal sheets flatness. Section 3 describes our noise model for the generation of synthetic data. Section 4 reviews computational approaches for image denoising, setting the stage for our proposal. Sections 5 and 6 present the proposed deep learning architecture for range image denoising and the collected Dataset, respectively. Section 7 reports the experimental results. Finally, Section 8 gives our conclusions and directions of future work.

2. Industrial Context

In order to inspect rolled products achieving accurate measurements and classification of flatness defects, it is necessary to capture the geometry of the steel sheet as it moves through the processing line. With sheet feeding rates reaching speeds of up to 120 m/min, real-time inspection imposes very strict requirements for accurate surface flatness quality control. The most typical flatness defects are wavy edges, centre buckles, and bow defects, which appear as low-frequency variations in the metal strip surface height.

On account of the strict requirements for real time quality control of surface flatness, the time efficiency of noise filtering methods poses a major challenge. Most of the literature [7,9,11] addresses this problem relying on the use of traditional filtering methods or explicit noise modeling, requiring extensive fine-tuning to adequately adapt to different noise levels, struggling in preserving details, and leading to local (sensor-specific) solutions.

Several successful applications of machine learning and fuzzy systems modeling for the detection of surface defects in flat steel products can be found in the literature [13–15], but they do not extend to the categorization of flatness defects. There are even machine learning approaches to link different types of defects with their causes [16,17].

Contrary to traditional hand crafted filtering methods, Convolutional Neural Networks (CNNs) are tuned by automated learning techniques guided by error minimization carried out by stochastic gradient descent and backpropagation algorithm. They have improved sensor data interpretation, analysis and control algorithms, being capable of dealing with non-linearities, noise, and uncertainty. In this regard, CNNs have become the state-of-the-art machine learning approach in many applications [18–22]. Recently, CNNs have been applied to classify surface defects in cold-rolled strips [23], and flatness measure prediction [24] from measurements of contact sensors attached to the roll mill instead of optical or range images of the surface. In order to adapt their 1D data from the sensor readings they fold these vectors into small images (5×8 or 20×20) which are the input for the CNNs, following the convention that CNNs are image classifiers or regressors. Note that the goal in [24] is the prediction of an overall measure of flatness from linear sensor readings.

However, to the best of the authors knowledge, there are no studies yet on CNN or other deep-learning-based methods to filter data obtained from optical flatness sensors in order to accurately reconstruct the surface of metal strips. In this regard, we are specifically interested in assessing the denoising performance of deep learning architectures when the input range image data contain high levels of non-linear noise.

Actual Sensor Installation

The flatness data were acquired with a simplified version of the optical flatness sensor described in [10]. The flatness sensor is comprised of a single illuminating linear laser source perpendicular to the metal sheet translation axis and a CCD camera capturing the area illuminated by the laser. In this simplified sensor version, the baseline separation between camera and the laser source is $\Delta_B = 900$ mm, and the triangulation angle is $\alpha = 45°$ so that the center of the camera captures the middle of the laser line at $Z = 0$ mm. The laser line emitter is collimated, and its wavelength is $\lambda = 450$ nm, while its line aperture is $90°$. The camera features a 2048×2048 matrix CCD sensor, and the focal length of the lens is $f = 6$ mm, placed at $Z = 1140$ mm over a moving steel strip. Figure 1 shows the scheme of the sensor.
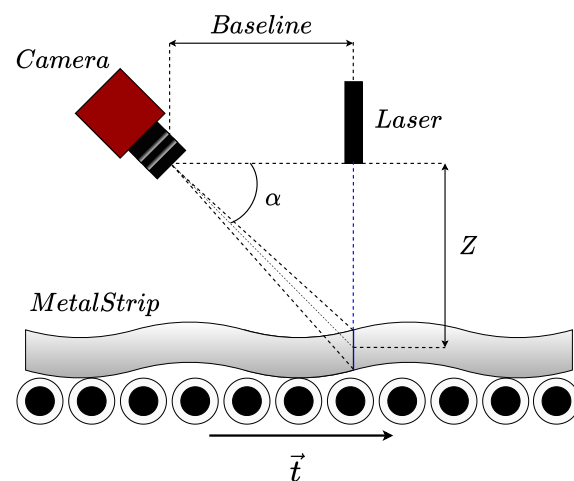
Figure 1. The flatness sensor scheme used for data acquisition, consisting of a single laser line laser triangulation scheme.

Figure 2 shows the scheme of the production line and the placement of the optical flatness sensor. Steel coils which are reduced to a specific thicknesses by rolling and

annealing and wound into a roll. These steel coils are further processed in a roll leveling and shearing line where they are cut to length. The range sensor was placed before the cutting tool, so the steel sheet surface propagates the vibrations induced by the cutting shocks. Each type of steel coil possesses different mechanical properties and thickness. As a result, they exhibit different propagation responses to the vibrations induced in the metal sheet during the leveling and cutting processes. This fact adds variability and robustness requirements to the proposed network.
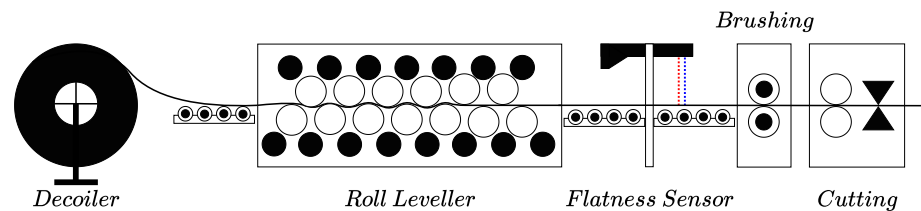
Figure 2. The experimental production line scheme and the optical flatness sensor placement. Blue and Red lines refer to laser planes used for pseudo-groundtruth calculation in real data experiments. Blue line refers to the laser plane which is further used for training, validation, and testing the proposed *CBRDNet*.

3. Noise Model for Synthetic Data Generation

Generating physically consistent surface data are crucial to train the proposed CBRD-Net and increase its denoising generalization capability. However, modeling such metal surfaces is impeded by the lack of accurate experimental data. Custom metrology devices, such as coordinate measuring machines (CMM), rely on static measuring conditions and, thus, fail to retrieve the most characteristic surface deformation caused by the tensile and trachle stresses occurring at the metal strip roll leveling and cut to length processes. To cope with this lack of data, our synthetic samples rely on a model of experimentally reconstructed surface data shown in [10], which reproduce the most common defects in a roll leveler processing line, as well as the coupling noise produced by mechanical elements, such as cutting stage.

We model the range image captured by our sensor from metal surface data by a function that combines a high-frequency and high-amplitude bump produced by the cutting stage, modeled as a local Gaussian signal $\psi(x, y)$, a superposition of a set of stationary waves $\varphi(x, y)$, a low-frequency carrier $\theta(y)$ and a Gaussian noise term $\rho(x, y)$ modeling the data acquisition electronics error,

$$S(x, y) = \varphi(x, y) + \psi(x, y) + \theta(y) + \rho(x, y) \tag{1}$$

where

$$
\begin{aligned}
\varphi(x, y) = {} & \sum_{n=0}^{N} \sum_{m=0}^{M} \alpha_{n,m} \cos\left(\frac{2\pi n x}{\lambda_x}\right) \cos\left(\frac{2\pi m y}{\lambda_y}\right) \\
& + \sum_{n=0}^{N} \sum_{m=0}^{M} \beta_{n,m} \cos\left(\frac{2\pi n x}{\lambda_x}\right) \sin\left(\frac{2\pi m y}{\lambda_y}\right) \\
& + \sum_{n=0}^{N} \sum_{m=0}^{M} \gamma_{n,m} \sin\left(\frac{2\pi n x}{\lambda_x}\right) \cos\left(\frac{2\pi m y}{\lambda_y}\right) \\
& + \sum_{n=0}^{N} \sum_{m=0}^{M} \delta_{n,m} \sin\left(\frac{2\pi n x}{\lambda_x}\right) \sin\left(\frac{2\pi m y}{\lambda_y}\right)
\end{aligned}
\tag{2}
$$

is a real-valued 2D Fourier series, where the amplitudes $\alpha = \delta = [0,5]$ and $\beta = \gamma = 0$. $\lambda_x = \lambda_y = [0, 0.1]$ are the wavelengths in the x and y directions,

$$\psi(x,y) = \frac{sin(f_b y A_b)}{1 + 4\left[\frac{(y-y_0)}{L_b}\right]^2} \tag{3}$$

is a high-frequency, high-amplitude Gaussian wave mixed with a low-frequency carrier modeling the bump produced by the cutting device, where $f_b = 5$ represents the bump carrier frequency, $A_b = [1,3]$ stands for the bump amplitude, $L_b = [10, 20]$ is the bump wave attenuation, and

$$\theta(y) = A_c cos(K_c y) \tag{4}$$

is a low-frequency carrier that sets the offset of the surface data along the transversal y-direction, where $A_c = [0, 0.5]$ is the carrier amplitude and $K_c = [0, 0.1]$ represents the frequency in the y direction. Finally, $\rho(x,y)$ is the electronic noise that arises during data acquisition caused by the discrete nature of radiation, i.e., the fact that the optical sensor captures an image by collecting photons. Considering some assumptions, this noise can be approximated by an additive model in which the noise has a zero-mean Gaussian distribution determined by its variance $\sigma_n^2 = [0.1, 0.35]$. That is, each value in the noisy data is the sum of the real value and a random, Gaussian distributed noise value. The defined intervals of variation and constant values for these variables have been selected in order to obtain synthetic data that are as close as possible to that acquired by the sensor in real experiments. We disregarded strict boundary conditions, such as Dirichlet conditions due to the free form nature of the unrolled metal coils on the machine. A synthetic surface generated using this model is shown in Figure 3.
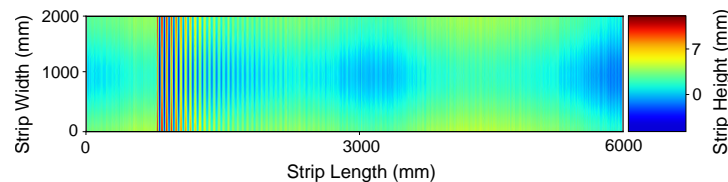
Figure 3. Synthetic flatness sensor data. (Color Online).

As shown in Figure 3, the proposed noise model allows us to generate synthetic data that are very similar to that acquired by the sensor in real experiments. The degree of concordance between our model and experimental data have been qualitatively validated by visual inspection. We cannot tune the model quantitatively because the noise source is not observable. We cannot observe the noise separated from the actual metal sheet surface, and the wave propagation and dumping properties are dependent of the actual metal sheet mechanical properties. We postulate that the success of the denoising system trained on the synthetic data are indirect proof of the validity of the model.

4. Deep Learning Denoising Approaches

An autoencoder is an unsupervised neural network architecture that is trained to reproduce the input as its output. It has a typical structure as a pair of funnels attached by the short end. The first funnel compresses the input data into a lower-dimension encoding, while the second funnel decompresses the encoding trying to recover the original input data. The encoder seeks to obtain a robust latent representation of the original data, which is often used for other purposes, such as features for another classification module. Autoencoders have been a popular field of study in neural networks in recent decades. The first applications of this type of neural networks date back to the 1980s [25–27]. Autoencoders have been used for classification, clustering, anomaly detection, dimensionality reduction, and signal denoising [28].

Proposed by Vincent et al. [29], the Denoising Autoencoders (DAEs) are an extension of classic autoencoders where the model is taught to predict original uncorrupted data from corrupted input data, i.e., the decoder attempts to reconstruct a clean version of the corrupted input from the autoencoder latent representation.

The encoder function f takes an input $\tilde{\mathbf{x}}$ and maps it to a hidden representation $\mathbf{y}$ computed as:

$$\mathbf{y} = f_\theta(\tilde{\mathbf{x}}) = h(\mathbf{W}\tilde{\mathbf{x}} + \mathbf{b}) \tag{5}$$

where h is a typically nonlinear transfer function, $\mathbf{W}$ and $\mathbf{b}$ are the encoder network parameters, and $\theta = (\mathbf{W}, \mathbf{b})$.

The output $\mathbf{x}$, having a similar form to $\tilde{\mathbf{x}}$, is reconstructed from $\mathbf{y}$ by the decoder g

$$\mathbf{x} = g_{\theta'}(\mathbf{y}) = h'(\mathbf{W'y} + \mathbf{b'}) \tag{6}$$

where h' is similar to h, $\mathbf{W'}$ and $\mathbf{b'}$ are the decoder network parameters, and $\theta' = (\mathbf{W'}, \mathbf{b'})$.

The DAE training procedure consists on learning the parameters $\mathbf{W}$, $\mathbf{W'}$, $\mathbf{b}$, and $\mathbf{b'}$ that minimise the autoencoder reconstruction error between the groundtruth $\mathbf{x}$ and the reconstruction $g_{\theta'}(f_\theta(\tilde{\mathbf{x}}))$, using a suitable cost function. Typically, the function is minimised using Stochastic Gradient Descent (SGD) [30] for small batches of corrupted and clean sample pairs.

Convolutional Denoising Autoencoders (CDAEs) are Denoising Autoencoders implemented using convolutional encoding and decoding layers. Because CDAEs use CNNs for extracting high-order features from images, CDAEs differ from standard DAEs in that their parameters are shared across all input image patches to maintain spatial locality. Different studies show that CDAEs achieve better image processing performance when compared to standard DAEs [31,32].

The U-Net [12] has a encoding–decoding architecture inspired in the autoencoder with skip connections [33] that transfer the data from the encoder layers to the decoding layers. Input–output pairs are images and their desired semantic pixel labelling providing segmentation of the image in one shot. It has shown exceptional results for image segmentation and image restoration tasks [34–36]. Depending on the architectural modifications made to U-Net, it can be used to achieve different tasks beyond segmentation. Isola et al. [37] used U-Net as a generator to perform image-to-image translation tasks such as in the case of aerial images and their correspondence in maps or the conversion of gray-scale images to color images through adversarial learning. Jansson et al. [38] investigated the use of U-Net as a voice separator, using the magnitude of the spectrogram of the audio containing the mix of different singing voices as the input. Zhang et al. [39] modified U-Net with a residual block and proposed it as a tool for extracting roads from aerial maps.

State-of-the-art 2D deep learning image denoising methods that will be compared with our proposal are CBDNet [40], NERNet [41], BRDNet [42], FFDNet [43], and CD-nCNN_B [44]. CBDNet is a convolutional blind denoising network [40] that is composed of a noise estimation module and a non-blind denoising module that accepts the noise estimation to compute the clean image. The noise estimation module is a CNN without pooling (i.e., no dimension reduction), while the denoising module is a U-shaped network as discussed above. The work reported in [40] uses a realistic noise model that includes in-camera processing to generate synthetic images with known noise component for network training. The noise estimation and removal network NERNet [41] inherits the two module structure of CBDNet. The noise estimation module is enriched with a pyramidal feature fusion block that provides multi-scale noise estimation, while the CNN components are dilated convolutions. The noise removal module is U-shaped using dense convolution and dilation selective blocks. The synthetic images were generated adding white Gaussian noise (AWGN). In the batch renormalization denoising network BRDNet [42], the batch renormalization is claimed to address the internal covariate shift and small mini-batch problems. The network is composed of upper and lower networks. Upper network is composed of residual learning modules with batch renormalization, while the lower net-

work includes also dilated convolution blocks. Contrary to the previous networks, no explicit noise estimation module is designed. Noise is assumed to be AWGN. The fast and flexible denoising network FFDNet [43] is also designed for cleaning AWGN corrupted images. FFDNet is a CNN whose inputs are downsampled subimages and a noise level map, it does not have a module to estimate the noise. The denoising convolutional neural network (DnCNNs) [44] is able to handle Gaussian denoising with unknown noise level. The DnCNN uses residual learning in order to estimate the noise component of the image, which is later removed from the noisy image to obtain the clean image.

## 5. Proposed Deep Learning Image Denoising Architecture

We apply of U-Net architecture as a generalized denoising method for surface reconstruction from noisy range images. The proposed network should be capable of denoising the degraded range images as an alternative to traditional image denoising techniques like spatial filtering, transform domain filtering, or wavelet thresholding methods [45]. A denoising method should remove high- and low-frequency noises, reconstructing the original surface. Results presented in the literature show that CNNs outperform traditional techniques for denoising tasks [46,47]. Furthermore, once trained, CNNs are computationally very efficient as they may be run on high-performance graphic processing units (GPUs) [48,49].

Our study proposes a convolutional blind residual denoising network model (CBRDNet) based on the U-Net architecture for denoising flatness sensor data. Since in real-world scenarios only noisy input data are provided, correct estimation of the noise level has proven to be challenging [40]. Therefore, incorporating a noise estimation block, can enhance the network generalization capabilities as shown by Lan et al. [50] and Guo et al. [41]. Besides that, the combination of both synthetic and real noisy data in the model training is expected to improve the network's denoising efficiency [51].

The structure and denoising functionality of the proposed network are described within the following sub-section.

5.1. Network Architecture

The proposed CBRDNet architecture consists of mainly two stages, a blind residual noise estimation subnetwork (NE-SNet) and a noise removal subnetwork (NR-SNet). The overall scheme of the proposed network is shown in Figure 4
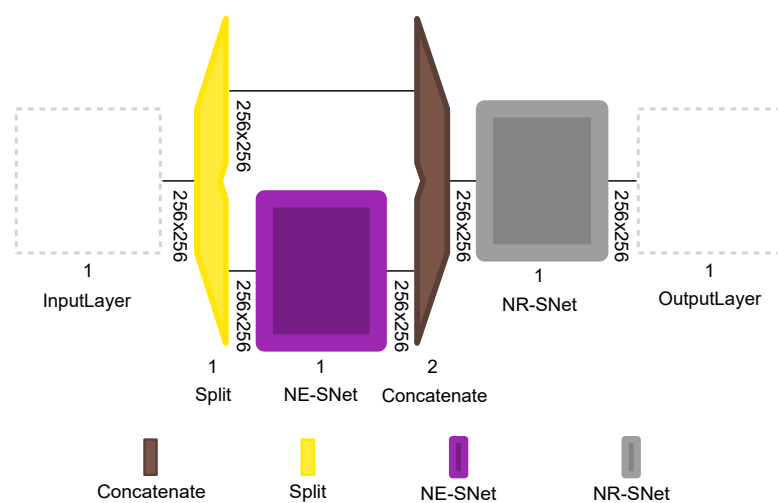
Figure 4. Overall scheme of the proposed CBRDNet network for close to real-time flatness data denoising.

The NE-SNet subnetwork takes a noisy observation and produces an estimated noise level map. It is composed of residual learning blocks that were first proposed as part of the ResNet architecture [52]. The layers of this subnetwork will increasingly separate image

structure from noise, creating a noise map that will be used later in the denoising stage. The NE-SNet is composed of five residual blocks with no pooling, each of which has two convolutional (Conv2D) layers with Batch Normalization (BN) and Rectified Linear Unit (ReLU) layers. The number of feature channels in each Conv2D layer is set to 64, and the filter size is set to 3×3. The scheme of the NE-SNet subnetwork is shown in Figure 5.
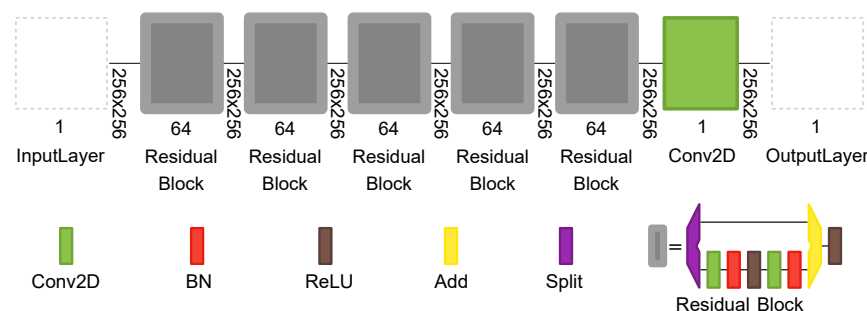
Figure 5. The proposed Noise Estimation Subnetwork NE-SNet composed of residual learning blocks.

The NR-SNet subnetwork is based on a traditional U-Net. This subnetwork is divided into two major paths: contracting (encoder) and expanding (decoder). The contracting path is comprised of downsampling layers consisting of a MaxPooling2D layer and two Conv2D layers with a filter size of 3×3 and "same" padding configuration. Each contracting block halves the size of feature maps and doubles the number of feature channels, starting with 64 channels in the first stage and ending with 512 channels in the last. The bottleneck connects both the expanding path and the contracting path; herein, the data has been resized to $32 \times 32 \times 512$. Similarly, the expanding path also comprises four upsampling blocks, which are composed of two Conv2D layers followed by a Conv2D Transpose. Each expanding block doubles the size of feature maps and halves the number of feature channels. We used concatenation layers to merge the feature maps in the expanding path with the corresponding feature maps in the contracting path. The last layer is a 1×1 Conv2D. The original U-Net architecture for image segmentation uses a sigmoid activation function in this last layer. Instead, our proposed architecture uses a linear activation function in order to recover the denoised image. The scheme of the NR-SNet subnetwork is given in Figure 6.
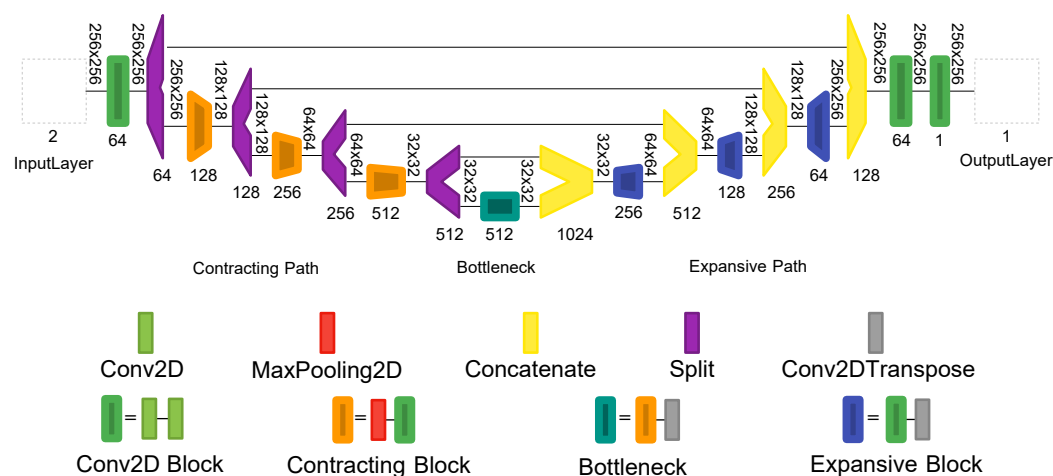
Figure 6. The proposed Noise Removal Subnetwork NR-SNet following a U-net architecture.

5.2. Training the Model

Given a 3D dataset encompassing data recovered from the laser based optical flatness sensor and synthetic 3D data described in Section 6, we generate a set of depth images,

which are decomposed into patches for processing. Using this dataset of local patches, we train our network to reconstruct the denoised versions of input depth images. In order to train the CBRDNet, we use the ADAM [53] algorithm with $\beta = 0.9$. Following most CNN-based data denoising methods, our network adopts the mean squared error (MSE) as the loss function and the initialization strategy of He [54]. The mini-batch size is 10, and each patch size is 256×256 pixels. The mini-batch size has been selected as a trade-off between our limited computational capabilities and the desired network generalization performance. Experimental results demonstrate that small batch sizes with small learning rates result in more reliable and stable training, better generalization performance, and a much lower memory footprint [55,56]. The model is trained for 100 epochs, with the learning rate for the first 20 epochs set to 10^{e-3} and the learning rate 10^{e-4} used to fine-tune the model. These settings are the same for all experiments discussed in this paper for uniformity. Besides that, both ReLU and LeakyReLu [57] have been tested as output layer activation functions in the CBRDNet training, the obtained results were almost identical and are shown in Section 7. We trained all the networks in this paper on a single NVIDIA® Geforce® RTX 2080 Super GPU with an on-board frame buffer memory of 8GB GDDR6, 3072 CUDA® Cores operating at 1815 MHz, compute capability 7.5, and Turing Generation microarchitecture, CUDA® 10.1 and CUDNN 7.6.1). The machine is equipped with an Intel® Core i9-10900K CPU @ 3.70GHz processor with 10 cores and 32 GB of RAM.

6. Dataset

The dataset used for both training and testing of the proposed architecture is composed of real production line and synthetic range image samples of steel coils from a roll levelling and shearing line. The synthetic data are used as a kind of data augmentation aiming to improve the network denoising performance because of the difficulties faced collecting a real dataset comprising a wide range of representative samples. Additionally, in real-world measurements the metal sheet is not free from tensile stresses during the manufacturing processes causing its elongation. After cutting the metal strip in single smaller sheets, the tensile stress release results in surface deformations. Thus, measurements obtained by an offline precision measuring device like a coordinate measuring machine (CMM) cannot be used as a validation ground truth for online measurement methods, whereas synthetic samples do.

In this paper, we generate 5500 synthetic noisy data samples using the noise model described in Section 3 together with 5500 real noisy samples from six different coils which are described in Section 6.1. The dataset is divided into a training set (80%), a validation set (10%) and a test set (10%).

### 6.1. Real Production Line Data

The experimental data coming from the real production line consists of 5500 samples from six different steel coils.

The specifications of the six steel coils are as follows: Two S235JR coils, a carbon (non-alloy) steel formulated for primary forming into wrought products with thicknesses of 3 mm, 8 mm and 1200 mm width, respectively, Young modulus $E = 205$ GPa, Poisson ratio $\mu = 0.301$, yield stress $\sigma = 215$ MPa, annealed and skin passed. One S420ML coil, a special structural steel with a thickness of 7 mm and 2000 mm width, Young modulus $E = 190$ GPa, Poisson ratio $\mu = 0.29$, yield stress $\sigma = 410$ MPa, it is an iron alloy steel manufactured by rolling. One S355M coil, an alloy steel formulated for primary forming into wrought products with a thickness of 3 mm and 1500 mm width, Young modulus $E = 190$ GPa, Poisson ratio $\mu = 0.29$, yield stress $\sigma = 360$ MPa, a middle carbon steel manufactured by rolling, annealing and skin passing. Two S500MC coils, a hot-rolled, high-strength low-alloy (HSLA) with excellent engineering bending and cutting characteristics with a thickness of 3 mm, 6 mm and 2200 mm width, respectively, Young modulus $E = 210$ GPa, Poisson ratio $\mu = 0.304$, yield stress $\sigma = 500$ MPa, produced through thermomechanical rolling. A summary is given in Table 1.

Table 1. Steel coils used for collecting real data and their specifications.

Coil	w × h (mm)	Young (GPa)	Poisson	Yield Stress (MPa)
S235JR	1050 × 3	205	0.301	215
S235JR	2000 × 8	205	0.301	215
S420ML	1650 × 7	190	0.290	410
S355M	1500 × 3	190	0.290	360
S500MC	1050 × 3	210	0.304	500
S500MC	1850 × 6	210	0.304	500

The coils are roughly 800 m long. In each measurement cycle, the optical flatness system senses 9000 mm. High-amplitude disruptive noises from the cutting station, as well as the mechanical processes carried out during the manufacturing greatly contaminate the flatness information generating noisy ripples on the metal strip sensor data. Additionally, the conveyor system generates high-frequency waves as a result of the metal strip advance. This interference patterns result in a complex spatial waveform, causing flatness information and surface defects difficult to detect. A raw depth data sample from one of these steel coils, captured by the optical flatness sensor, is visualised in Figure 7.
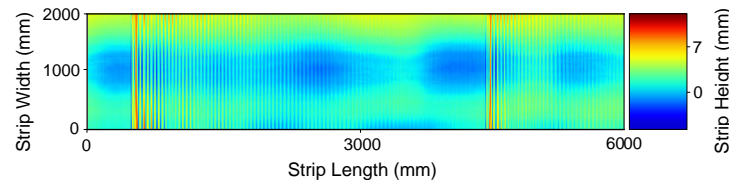
**Figure 7.** Raw optical flatness sensor data. (Color online).

7. Results

In this section, we assess the proposed CBRDNet for denoising both synthetic sheet samples and real data from the 3D flatness sensor. The proposed denoising network is employed to reconstruct both simulated and real data in order to test its ability to remove non-linear noises caused by mechanical manipulation of the metal sheet during the manufacturing process.

The metal sheet's flatness corresponds to its levelness when it is tension free. The I-Unit [58] is widely used as the standardized measurement unit of flatness. For the I-Unit calculation in a metal sheet with a sinusoidal surface, a series of virtual lines are drawn to model the surface profile. The I-Unit is computed over them and the reported flatness is the average over all lines. For this reason we compare our 2D methods with 1D denoising methods. We recall that the aim of the present work is to provide a CNN-based denoising method to be be applied to range images obtained by optical sensors installed in metal sheet leveling and shearing production lines. The denoised surface range data will be used to carry out the necessary flatness measurement. Accordingly, the results provided below compare the denoised synthetic sheet samples and real ones with its corresponding groundtruth. The error measurements are expressed in millimeters.

7.1. Synthetic Data Results

We conducted three different comparative analyses. First, we apply some traditional 1D filtering methods such as Moving Average, Butterworth IIR [59,60], Savitzky-Golay FIR [61,62], Chebyshev Type II [63], and piecewise cubic Hermite interpolation [10] filters. Secondly, we apply 2D wavelet-based denoising methods. Specifically, we compute results using Daubechies, Symlets, Meyer, Coiflets, and Fejer-Korovkin wavelets [64–66]. Finally, we compare the performance of CBRDNet against some state-of-art 2D deep learning image denosing methods, specifically NERNet [41], CBDNet [40], BRDNet [42], FFDNet [43], and CDnCNN_B [44]. Instances of synthetic data denoising results are shown in Figures 8 and 9, where (a) is the noise-free sample, (b) is the noisy surface data and, finally, (c) is the denoised surface estimated using our method.
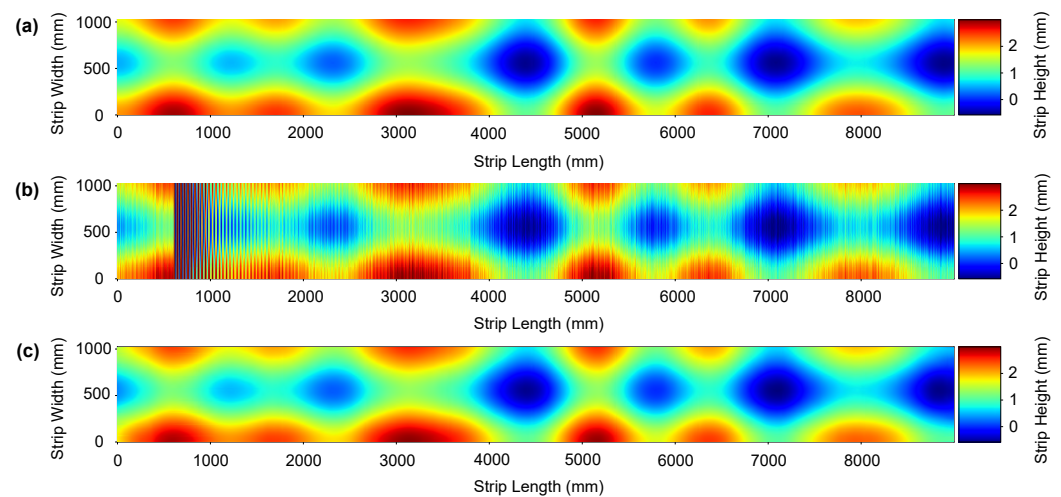
Figure 8. An instance of the denoising result on a synthetic strip. (**a**) Depicts the noise-free ground truth surface, (**b**) shows the noise corrupted surface, and (**c**) represents the denoised surface reconstructed using the proposed network. (Color online).
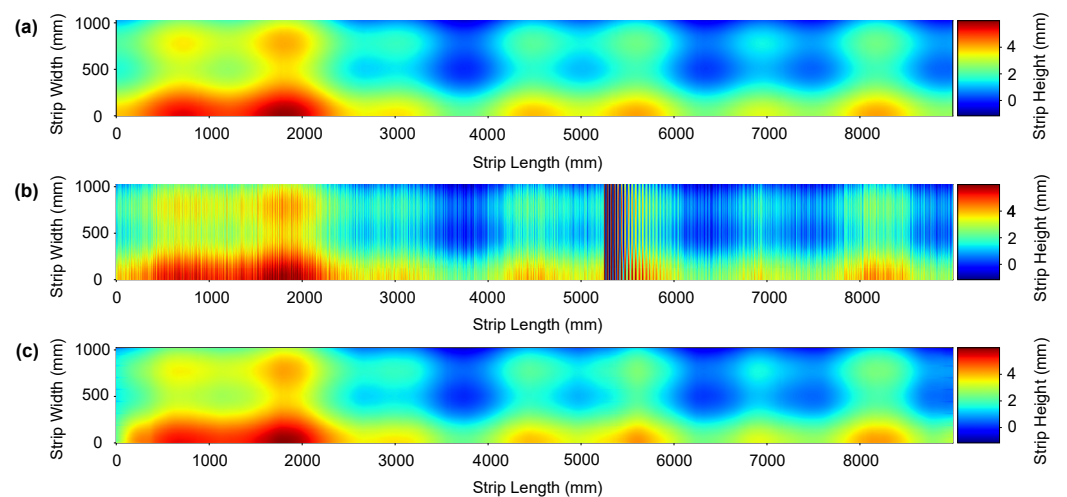
Figure 9. An instance of the denoising result on a synthetic strip. (**a**) Depicts the noise-free ground truth surface, (**b**) shows the noise corrupted surface, and (**c**) represents the denoised surface reconstructed using the proposed network. (Color online).

For the comparative analysis with traditional 1D filtering methods we divided the resulting metal sheet surface in virtual longitudinal strips, also called fibers [58,67]. For each fiber, we applied the following 1D denoising approaches:

A Butterworth IIR filter. This filter provides the optimum balance of attenuation and phase response. It has no rippling effect in the passband or stopband, and as a result, it is frequently referred to as a maximally flat filter. The Butterworth filter provides flatness at the cost of a somewhat broad transition area from passband to stopband, with typical transitory characteristics. It has the following characteristics: a smooth monotonic response (no ripple), it has the slowest roll-off for equivalent order filters, and a more linear passband phase response than other methods. A Butterworth IIR third-order digital filter with a cutoff frequency of 6 dB below the passband value of 0.01 specified in normalized frequency units is used.

A Savitzky-Golay FIR smoothing filter, which is a variation of the FIR average filter that can effectively retain the targeted signal's high-frequency content while still not eliminating as much noise as a FIR average. Savitzky-Golay filters maintain various moment orders better than other smoothing approaches, which generally retain peak widths and heights. It has the following characteristics: a computation time proportional to window width,

it preserves the area, position and width of peaks, and flattens peaks less than moving average with same window width. A third-order Savitzky-Golay FIR smoothing filter with a frame length of 99 samples is used in our experiments.

A Moving Average filter was also applied, which is a method used to smooth data by calculating a series of averages of different subsets of the entire dataset. It is a form of finite impulse response filter with the following characteristics: an optimal approach for reducing random noise while retaining a sharp step response, in general term is a good smoother filter, conceptually it is the simplest to implement, but on the contrary has a poor low-pass filter (frequency domain) and a slow roll-off and terrible stopband attenuation characteristics. A moving-average filter with a 33-sample-long sliding window is used for the comparison experiments.

A Chebyshev Type II filter has been applied. This filter is also known as an inverse filter, it does not roll off and has no ripple in the passband, but it has equiripple in the stopband. The main characteristics of this filter are: it is maximally flat in the passband and has a faster roll-off than Butterworth but slower roll-off than Chebyshev Type I. We used a third-order low-pass Chebyshev Type II filter with a stopband attenuation of 33 dB and a stopband edge frequency of 0.02 specified in normalised frequency units.

Finally, a piecewise cubic Hermite interpolation filter has been used. This filter uses both the height surface information and its derivative calculated from a dual laser sensor data series. It is continuous in shape and its derivative. In comparison to the Savitzky–Golay, Butterworth, Chebyshev, and Average Mean filters used for surface reconstruction in [10], this method achieved a 41 percent improvement.

Because we have the ground truth surface, we can compute the error of our denoising process. Table 2 shows the comparative results of the denoising approaches described above when applied to the synthetic surface. MAE improvements achieved by our method range from three times better when compared to the Hermite filtering approach to 6 times better when compared to the Chebyshev filter approach. Similar improvements are achieved in term of RMSE.

In addition, we conducted 2D wavelet-based denoising methods. The number of vanishing moments N and the denoising threshold are the metaparameters for this approach. According to the current research, disregarding the computational cost of the wavelet transform (WT), higher vanishing moments would yield better performance [68,69]. We selected the following wavelets: Daubechies (*dbN*), $N = 4$, Symlets (*symN*) $N = 8$, Meyer (*dmey*), Coiflets (*coif N*), $N = 4$, and Fejer-Korovkin (*f kN*), $N = 4$. We performed the WT of data samples up to 8 levels. For denoising, wavelet transform coefficients below an empirically selected WT threshold are set to zero. An inverse wavelet transform is used after that to transform the processed signal back to the original spatial domain. Because the wavelet coefficients are affected by values outside the extent of the signal under consideration, to avoid boundary effects, the first and last 4 samples were removed in the processed input data. Table 2 shows the comparative results. MAE improvements achieved by our method range from 2.5 times better when compared to the Fejer-Korovkin filtering approach to 1.3 times better when compared to the Symlets filter approach. Similar improvements are achieved in term of RMSE. For a graphical representation of these results, we provide the denoising results on five data samples in Figure 10.

Table 2. Comparative results of our approach with both traditional 1D and 2D denoising approaches and 2D denoising CNN methods. MAE = mean absolute error; MaxAE = maximum absolute error; STD = standard deviation of the absolute error; RMSE = root mean squared error. Best results presented in bold font.

| Method	CNN-2D/1D/2D	Blind/Non Blind	MAE *	MaxAE *	STD *	RMSE *
CBRDNet-ReLu (ours)	CNN-2D	Blind	**0.140**	**0.376**	**0.136**	**0.147**
CBRDNet-LeakyReLu (ours)	CNN-2D	Blind	0.160	0.466	0.154	0.172
CBDNet	CNN-2D	Blind	0.172	0.520	0.162	0.185
NERNet	CNN-2D	Blind	0.184	0.499	0.175	0.195
BRDNet	CNN-2D	Blind	0.198	0.659	0.184	0.212
FFDNet	CNN-2D	Non Blind	0.224	0.501	0.201	0.252
CDnCNN_B	CNN-2D	Blind	0.312	0.840	0.308	0.342
Sym8	2D	NA	0.176	0.543	0.170	0.188
Coif4	2D	NA	0.180	0.591	0.179	0.190
Db8	2D	NA	0.181	0.622	0.179	0.201
Dmey	2D	NA	0.256	0.942	0.282	0.291
Fk8	2D	NA	0.390	1.998	0.588	0.390
Hermite	1D	NA	0.413	1.150	0.380	0.459
Butterworth	1D	NA	0.760	4.423	0.735	0.781
Savitzky-Golay	1D	NA	0.842	6.436	0.779	0.853
Moving Average	1D	NA	0.801	5.463	0.928	0.865
Chebyshev Type II	1D	NA	0.828	5.040	0.828	0.903

* Measurements are expressed in millimeters (mm).
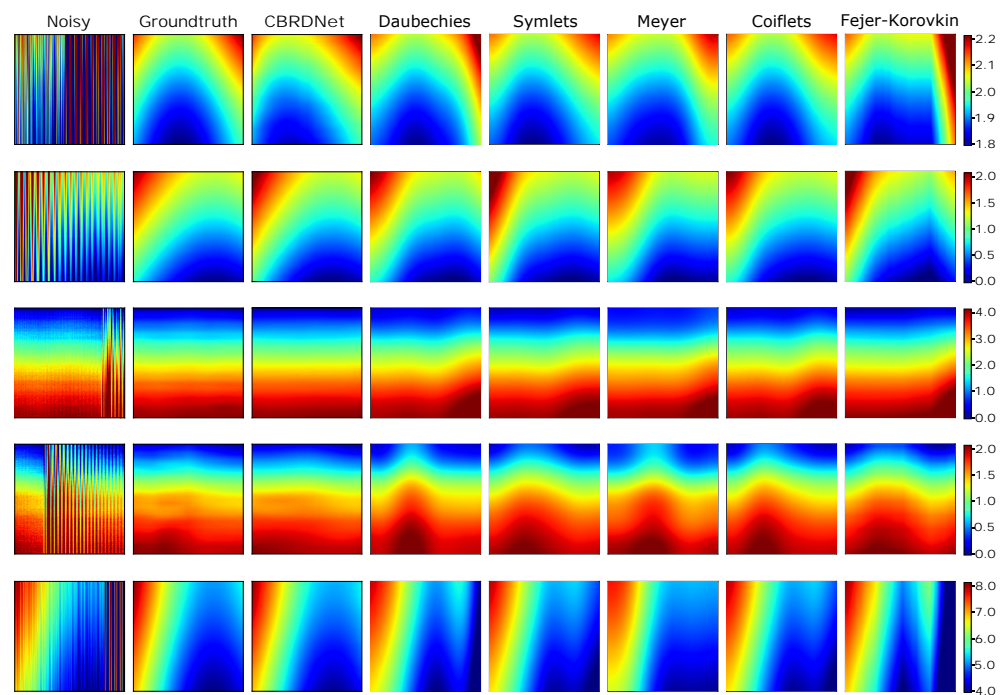
Figure 10. Comparison of our proposed CBRDNet with 2D wavelet-based denoising techniques. Noisy synthetic samples have very low SNR, hence the the groundtruth surface of the samples is almost lostt in some samples. To facilitate the comparison with the denoised samples, the color scale of the images corresponding to the first column, i.e., noisy sample, is clipped. Color scale values are expressed in millimeters (mm). (Color online).

Finally, we compared the architecture presented in this article to the five earlier stated CNN-based approaches. Comparing various deep learning algorithms is a challenging task because of the large number of hyperparameters that must be appropriately tuned during the network training process. Notwithstanding, the aforementioned architectures were trained and assessed 100 times on the same dataset to obtain the statistical results listed

in Table 2. Furthermore, for a clearer graphical representation of denoising performance, we provide the outcomes of these methods on five data samples, see Figure 11. When compared to the groundtruth the CBRDNet results are very close to the real ones, MAE improvements range from 2.5 times better when compared to CDnCNN_B and 1.2 times better when compared to CBDNet. Similar improvements are measured in terms of RMSE.
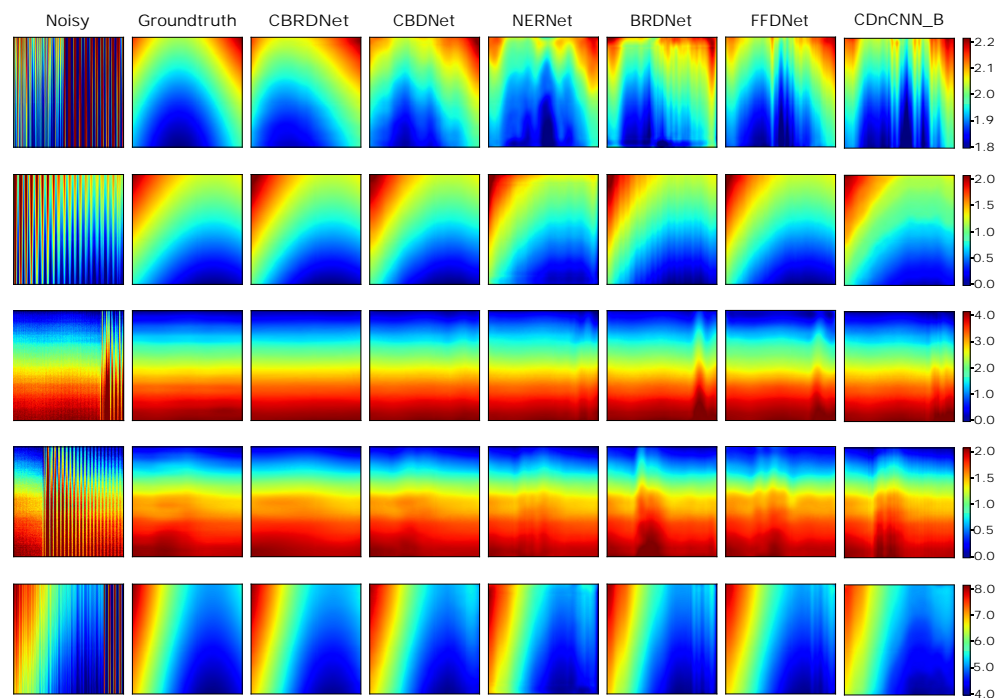

Figure 11. Comparison of our proposed CBRDNet with other methods on five different samples. Note that because of the low SNR, the geometrical surface of the samples is hidden by the induced perturbations. To facilitate the comparison with the denoised samples, the color scale of the images corresponding to the first column, i.e., noisy sample, is clipped. Color scale values are expressed in millimeters (mm). (Color online).

7.2. Real Data Results

Measuring results from a specimen tested out of the roll levelling system with a CMM cannot be fairly compared to those obtained by our method, as has been previously discussed in Section 6. Results obtained with the double laser line sensor and the Hermite filtering method proposed by Alonso et al. [10] have been used as groundtruth in order to evaluate the improvement of the proposed method in an industrial environment. Experimental results with real data are shown in Figures 12 and 13, where (a) is the denoised data using Hermite cubic interpolation, (b) is the raw data retrieved from the sensor and, finally, (c) is the denoised surface obtained using our method. The proposed CBRDNet architecture effectively recovers the smooth reconstructed surface after the noisy waves have been filtered, as seen in the figures.

The results shows graphically that the proposed method is capable of accurately reconstructing the surface of the metal sheet. When compared to state-of-the-art techniques, it achieves equivalent or better visually appealing results, as a real ground truth is always lacking in real experiments. Figure 14 depicts a longitudinal fibre, with unfiltered data collected directly from the sensor in blue, Hermite filtering in red, 2D Symlet wavelet-based filtering results in yellow, and the results from the CNN proposed in this work in green. It can be seen that the method is capable of reconstructing the sheet's surface preserving the sinusoidal characteristics of the metal sheet, specially in areas where the cutting effect occurs.
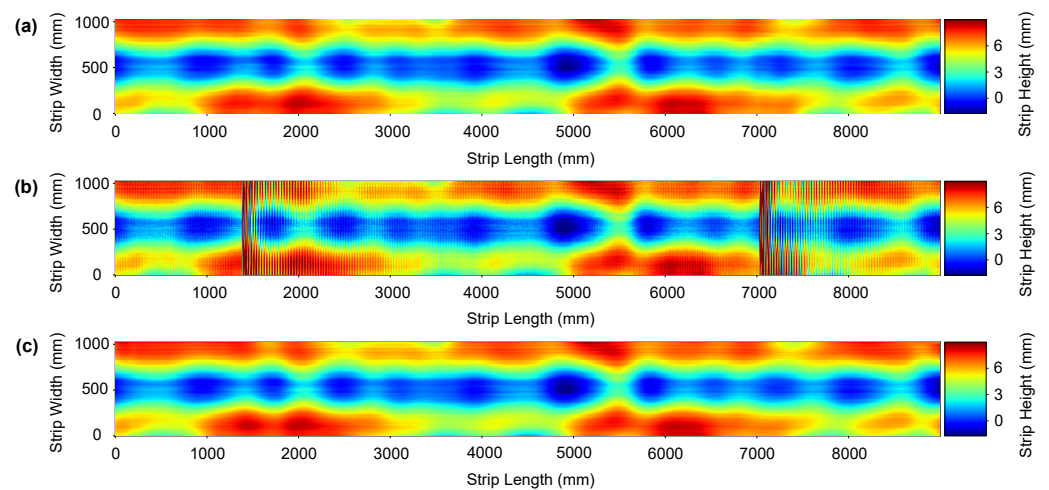
Figure 12. An instance of a real metal sheet surface denoising. Plate thickness: 3 mm; yield point Re: 215 MPa; Dimension of the mother plate—length: 9 m; width: 1050 mm. (**a**) Hermite filter denoised ground truth surface, (**b**) noise corrupted surface and (**c**) denoised reconstructed surface using the proposed network. (Color online).
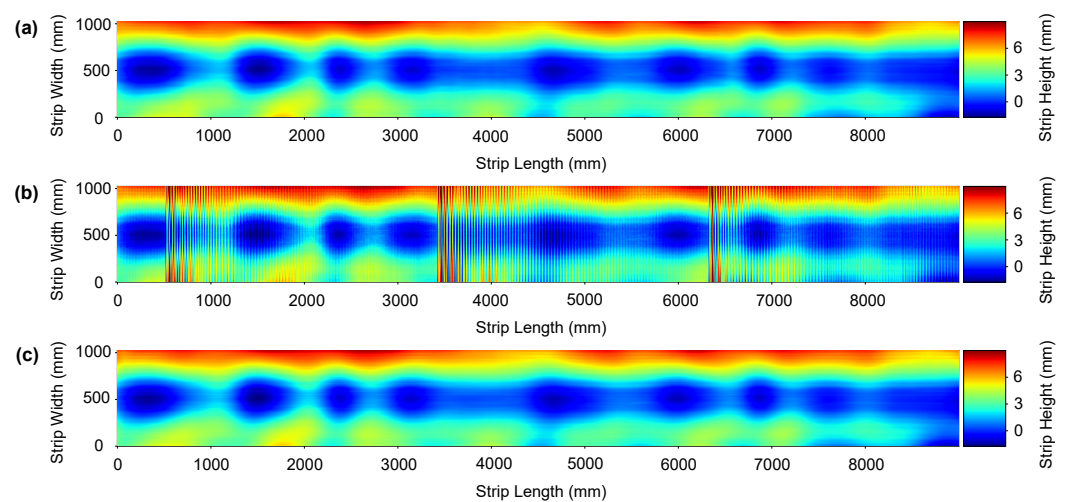
Figure 13. An instance of a real metal sheet surface denoising. Plate thickness: 6mm; yield point Re: 500 MPa; Dimension of the mother plate—length: 9 m; width: 1050 mm. (**a**) Hermite filter denoised ground truth surface, (**b**) noise corrupted surface and (**c**) denoised reconstructed surface using the proposed network. (Color online).
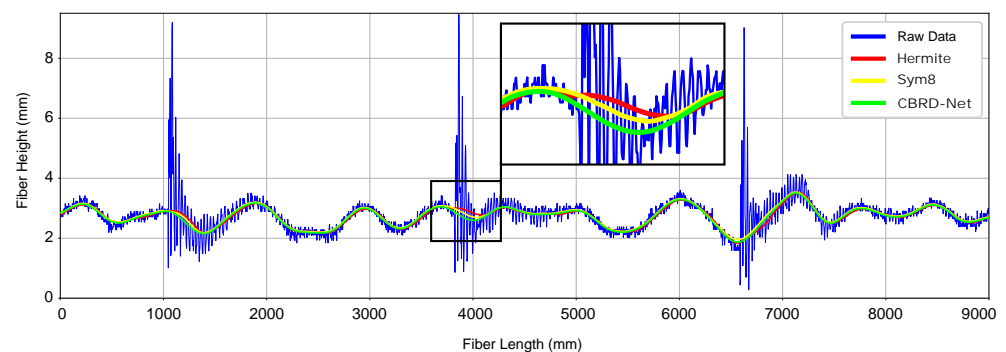
**Figure 14.** An instance of longitudinal fiber reconstruction, the blue line represents the raw 1D data from a fiber extracted from Figure 13, the red line shows the Hermite interpolation result, the yellow line shows Symlet results, and finally the green line depicts the CBRDNet denoised result. The inset provides a better detail of the results achieved by both Hermite and CBRDNet in the highlighted area.

7.3. Ablation Studies

Several ablation studies have been carried out in order to analyse the effects of both the noise estimation module (NE-SNet subnetowk) and training the network with synthetic, real, and mixed datasets.

7.3.1. Effect of the NE-SNet Subnetwork

An ablation study was conducted to better understand the contribution of the NE-SNet subnetwork component to the overall system. This research has revealed that the overall performance of the proposed system is highly dependant on the NE-SNet subnetwork, increasing the accuracy of the proposed network up to 10%. Quantitative results of this study are shown in Table 3. Besides that, noise prediction experiments reveal that the NE-SNet achieves an accuracy of nearly a 85% extracting the noise both in synthetic and real data. Figure 15 depicts some results obtain by the NE-SNet subnetwok over both synthetic and real metal strip patches. The mean absolute error (MAE), maximum absolute error (MaxAE), standard deviation of the absolute error (STD), and root mean squared error (RMSE) were evaluated over a 500 sample dataset, results are as follows, MAE = 0.420 mm, MaxAE = 1.105 mm, STD = 0.124, and RMSE = 0.480 mm.
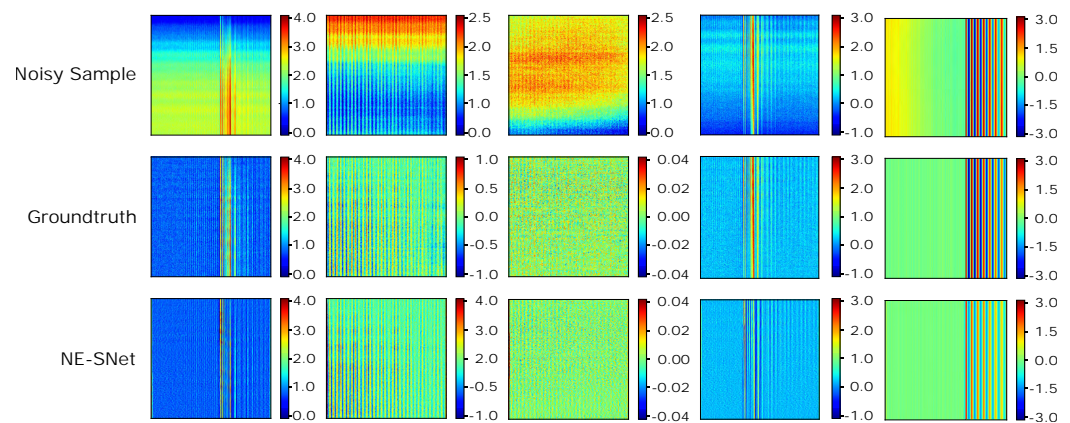
Figure 15. Results of the proposed NE-SNet subnetwork estimating the noise level map over a set of real and synthetic data samples. Noisy Sample stands for the corrupted noisy data, groundtruth is the known noise level map, finally NE-SNet is the estimated noise output. Color scale values are expressed in millimeters (mm). (Color online).

Table 3. Comparative results of our NE-SNet subnetwork ablation study with the full model and the best CNN and conventional denoising approaches. MAE = mean absolute error; MaxAE = maximum absolute error; STD = standard deviation of the absolute error; RMSE = root mean squared error. Best results presented in bold font.

Method	CNN-2D/1D/2D	MAE *	MaxAE *	STD *	RMSE *
CBRDNet (Full Model)	CNN-2D	**0.140**	**0.376**	**0.136**	**0.147**
CBRDNet (No NE-SNet)	CNN-2D	0.305	1.043	0.284	0.385
CBDNet	CNN-2D	0.172	0.520	0.162	0.185
Sym8	2D	0.176	0.543	0.170	0.188
Hermite	1D	0.413	1.150	0.380	0.459

* Measurements are expressed in millimeters (mm).

7.3.2. Effect of Synthetic and Real Data

We have developed the following approaches. First, we trained our proposed CBRD-Net on synthetic data exclusively. Second we trained CBRDNet on real data only. On the one hand, the experiments carried out demonstrate that CBRDNet (Synth) achieve worse results than CBRDNet (Real) and CBRDNet removing the existing real noise. This fact occurs even when trained on large amount of synthetic data samples, mainly because real noise cannot be accurately described by the defined noise model at 3. On the other hand, CBRDNet (Real) produces not so accurate results in comparison to CBRDNet, as a result of the impact of insufficiently noise-free real data. At the same time, CBRDNet has proved to be more effective in dealing with real noise while maintaining an accurate surface information. Quantitative results of the three strategies are shown in Table 4 on 500 sample synthetic, real, and mixed datasets. CBRDNet obtains better results than CBRDNet(Synth) and CBRDNet(Real) except in the synthetic dataset, but we dismiss these results as they are not directly applicable to a real production environment where real noise is present.

Table 4. Comparative results of training data ablation studies. MAE = mean absolute error; MaxAE = maximum absolute error; STD = standard deviation of the absolute error; RMSE = root mean. (Synth) = trained on synthetic dataset; (Real) = trained on real dataset squared error. Best results presented in bold font.

Method	MAE *	MaxAE *	STD *	RMSE *
Mixed dataset results				
CBRDNet	**0.140**	**0.376**	**0.136**	**0.147**
CBRDNet (Synth)	0.260	0.496	0.248	0.265
CBRDNet (Real)	0.180	0.401	0.175	0.186
Synthetic dataset results				
CBRDNet	0.190	0.410	0.181	0.195
CBRDNet (Synth)	**0.110**	**0.206**	**0.128**	**0.129**
CBRDNet (Real)	0.280	0.526	0.254	0.292
Real dataset results				
CBRDNet	**0.147**	**0.386**	**0.142**	**0.154**
CBRDNet (Synth)	0.282	0.366	0.265	0.291
CBRDNet (Real)	0.159	0.396	0.155	0.161

* Measurements are expressed in millimeters (mm).

8. Conclusions and Future Work

In this paper, we present a novel denoising deep learning architecture for filtering range image sensor data that can be used for accurate flatness measurement in the context of metal sheet manufacturing, named CBRDNet.

This network is able to filter out the non-linear noise components in the range images that hinder accurate surface reconstruction and thus surface flatness measurements. It has been trained using both real and synthetic samples of steel coils from a roll leveling cut to length line. This combination improves the network's denoising capabilities. Furthermore, synthetic data not only provided a wide range of representative samples for training, but also a groundtruth for quantitative evaluation of the accuracy of the denoised flatness measurements. We carried out different experiments to validate the proposed filtering strategy.

In the first place, results obtained denoising synthetic data have proved that our method outperforms traditional 1D filtering techniques, namely Hermite, Savitzky-Golay, Chebyshev, and Butterworth filters. Compared to them, we achieved an improvement of up to 6 times in terms of accuracy, particularly in surface regions where high amplitude noises are induced by the mechanical processes carried out in the production line, e.g., cutting the metal strip to the desired length. In the second place, the proposed CBRDNet achieves slightly better results in comparison with 2D wavelet-based filtering techniques.

We achieved an error reduction up to 1.3 times when compared to the best performing wavelet in our study, i.e., Symlets (Sym8), although in some sample regions there was no clear improvement in terms of precision. Wavelet denoising results must be taken with a grain of salt, because an optimal wavelet class and order selection might improve them, while we report results of a necessarily limited empirical exploration. To this date we do not know of such a data driven optimal wavelet design process. In the third place, experiments with synthetic data show that the CBRDNet architecture is able to obtain better results than state-of-the-art deep learning denoising architectures for the specific kind of noise that we are dealing with. Compared to these methods we obtain improvements ranging from 1.2 up to 2.5 times in terms of surface reconstruction accuracy. This improvement is clearly visible in the areas of the metal sheet where the noise due to metal strip cutting occurs.

Finally, results with real data obtained from an industrial leveling cut to length line have shown that the proposed method is capable of accurately reconstructing metal sheet surfaces. The conducted experiments have shown a surface reconstruction error reduction than can be down to 15% relative to solutions based on conventional interpolation methods. Numerical results have shown that the proposed CBRDNet achieves a mean absolute error (MAE) of 0.140mm a maximum absolute error (MaxAE) of 0.376 mm, a standard deviation of the absolute error (STD) of 0.136 mm, and a root mean squared error (RMSE) of 0.147 mm.

Future research will explore deep denoising architectures in the frequency domain. Although in some cases it is difficult to differentiate a signal from noise in the spatial domain, this task might be easier in the frequency domain because noisy signals can be comprised of a set of sine wave signals represented in the frequency domain with different frequencies, phases, and amplitudes. We intend to implement and compare these possible enhancements to the network outlined in this paper in future works. Moreover, when larger data sets are needed but the access to real data is restricted in some way, for example, when data becomes sensitive to its distribution, or simply when access to real data is challenging, the development of tools capable of generating synthetic data would provide a solution to this data shortage. GANs are computational structures that employ two neural networks, competing with each other, to create new synthetic data samples that may be used as surrogates for real data. To further our research we plan to explore the potential of using GANs architectures instead of the current noise model to generate larger dataset with more likelihood to real data.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
DAE	Denoising Autoencoders
CDAE	Convolutional Denoising Autoencoders
GAN	Generative Adversarial Network
SGD	Stochastic Gradient Descent
CBRDNet	Convolutional Blind Residual Denoising Network
NE-SNet	Noise Estimation Subnetwork
NR-SNet	Noise Removal Subnetwork
ADAM	Adaptive Moment Estimation
ReLU	Rectified Linear Unit
BN	Batch Normalization
Conv2D	2D Convolution Layer
HSLA	High-Strength Low-Alloy
MSE	Mean Squared Error
MAE	Mean Absolute Error
MaxAE	Maximum Absolute Error
STD	Standard Deviation
RMSE	Root Mean Squared Error
CMM	Coordinate Measuring Machine
Db	Daubechies
Coif	Coiflets
Sym	Symlets
Fk	Fejer-Korovkin
Dmey	Meyer

References

1. Jouet, J.; Francois, G.; Tourscher, G.; de Lamberterie, B. Automatic flatness control at Solmer hot strip mill using the Lasershape sensor. *Iron Steel Eng.* **1988**, *65*, 50–56.
2. Chiarella, M.; Pietrzak, K.A. An Accurate Calibration Technique for 3D Laser Stripe Sensors. In *Optics, Illumination, and Image Sensing for Machine Vision IV*; Svetkoff, D.J., Ed.; International Society for Optics and Photonics, SPIE, Bellingham: Washington, CA, USA, 1990; Volume 1194, pp. 176–185. [CrossRef]
3. Álvarez, H.; Alonso, M.; Sánchez, J.R.; Izaguirre, A. A Multi Camera and Multi Laser Calibration Method for 3D Reconstruction of Revolution Parts. *Sensors* **2021**, *21*, 765. [CrossRef]
4. Siekański, P.; Magda, K.; Malowany, K.; Rutkiewicz, J.; Styk, A.; Krzesłowski, J.; Kowaluk, T.; Zagórski, A. On-line laser triangulation scanner for wood logs surface geometry measurement. *Sensors* **2019**, *19*, 1074. [CrossRef] [PubMed]
5. Kierkegaard, P.; Classon, L.A. A new-generation optical flatness measurement system. *Iron Steel Technol.* **2015**, *12*, 76–81.
6. Kierkegaard, P. Developments and Benefits from Optical Flatness Measurement in Strip Processing Lines. In Proceedings of the Congreso y Exposición Nacional de la Industria del Acero, CONAC, Monterrey, Mexico, 8–20 November 2016; pp. 1–15.
7. Pernkopf, F. 3D surface acquisition and reconstruction for inspection of raw steel products. *Comput. Ind.* **2005**, *56*, 876–885. [CrossRef]
8. Usamentiaga, R.; Molleda, J.; Garcia, D.F.; Bulnes, F.G. Removing vibrations in 3D reconstruction using multiple laser stripes. *Opt. Lasers Eng.* **2014**, *53*, 51–59. [CrossRef]
9. Usamentiaga, R.; Garcia, D.F. Robust registration for removing vibrations in 3D reconstruction of web material. *Opt. Lasers Eng.* **2015**, *68*, 135–148. [CrossRef]
10. Alonso, M.; Izaguirre, A.; Andonegui, I.; Graña, M. Optical Dual Laser Based Sensor Denoising for Online Metal Sheet Flatness Measurement Using Hermite Interpolation. *Sensors* **2020**, *20*, 5441. [CrossRef]
11. Alonso, M.; Izaguirre, A.; Andonegui, I.; Graña, M. An Application of Laser Measurement to On-Line Metal Strip Flatness Measurement. In Proceedings of the 15th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2020), Burgos, Spain, 16–18 September 2020; Springer International Publishing: Cham, Switzerland, 2021; pp. 835–842.
12. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.
13. Borselli, A.; Colla, V.; Vannucci, M.; Veroli, M. A fuzzy inference system applied to defect detection in flat steel production. In Proceedings of the International Conference on Fuzzy Systems, Barcelona, Spain, 18–23 July 2010; pp. 1–6. [CrossRef]

14. Borselli, A.; Colla, V.; Vannucci, M. Surface Defects Classification in Steel Products: A Comparison between Different Artificial Intelligence-based Approaches. In Proceedings of the 11th IASTED International Conference on Artificial Intelligence and Applications, AIA 2011, Innsbruck, Austria, 14–16 February 2011. [CrossRef]

15. Xu, K.; Xu, Y.; Zhou, P.; Wang, L. Application of RNAMlet to surface defect identification of steels. *Opt. Lasers Eng.* **2018**, *105*, 110–117. [CrossRef]

16. Brandenburger, J.; Colla, V.; Nastasi, G.; Ferro, F.; Schirm, C.; Melcher, J. Big Data Solution for Quality Monitoring and Improvement on Flat Steel Production. *IFAC-PapersOnLine* **2016**, *49*, 55–60. [CrossRef]

17. Appio, M.; Ardesi, A.; Lugnan, A. Automatic Surface Inspection in Steel Products ensures Safe, Cost-Efficient and Timely Defect Detection in Production. In Proceedings of the AISTech-Iron and Steel Technology Conference, São Paulo, Brazil, 2–4 October 2018; pp. 89–101.

18. Graña, M.; Alonso, M.; Izaguirre, A. A Panoramic Survey on Grasping Research Trends and Topics. *Cybern. Syst.* **2019**, *50*, 40–57. [CrossRef]

19. Chong, E.; Han, C.; Park, F. Deep Learning Networks for Stock Market Analysis and Prediction: Methodology, Data Representations, and Case Studies. *Expert Syst. Appl.* **2017**, *83*, 187–205. [CrossRef]

20. Chen, X.W.; Lin, X. Big Data Deep Learning: Challenges and Perspectives. *IEEE Access* **2014**, *2*, 514–525. [CrossRef]

21. Najafabadi, M.M.; Villanustre, F.; Khoshgoftaar, T.M.; Seliya, N.; Wald, R.; Muharemagic, E. Deep learning applications and challenges in big data analytics. *J. Big Data* **2015**, *2*, 1. [CrossRef]

22. Affonso, C.; Rossi, A.L.D.; Vieira, F.H.A.; de Leon Ferreira de Carvalho, A.C.P. Deep learning for biological image classification. *Expert Syst. Appl.* **2017**, *85*, 114–122. [CrossRef]

23. Liu, Y.; Geng, J.; Su, Z.; Zhang, W.; Li, J. Real-Time Classification of Steel Strip Surface Defects Based on Deep CNNs. In *Proceedings of 2018 Chinese Intelligent Systems Conference, Vol II*; Springer: Singapore, 2019; pp. 257–266. [CrossRef]

24. Wang, Y.; Li, C.; Peng, L.; An, R.; Jin, X. Application of convolutional neural networks for prediction of strip flatness in tandem cold rolling process. *J. Manuf. Process.* **2021**, *68*, 512–522. [CrossRef]

25. Hinton, G.E.; Zemel, R.S. Autoencoders, minimum description length, and Helmholtz free energy. *Adv. Neural Inf. Process. Syst.* **1994**, *6*, 3–10.

26. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

27. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]

28. Bank, D.; Koenigstein, N.; Giryes, R. Autoencoders. *arXiv* **2021**, arXiv:2003.05991.

29. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.A. Extracting and Composing Robust Features with Denoising Autoencoders. In Proceedings of the 25th International Conference on Machine Learning, ICML'08, Helsinki, Finland, 5–9 July 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 1096–1103. [CrossRef]

30. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2017**, arXiv:1609.04747.

31. Gondara, L. Medical Image Denoising Using Convolutional Denoising Autoencoders. In Proceedings of the 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, Spain, 12–15 December 2016; IEEE Computer Society: Los Alamitos, CA, USA, 2016; pp. 241–246. [CrossRef]

32. Masci, J.; Meier, U.; Cireşan, D.; Schmidhuber, J. Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction. In Proceedings of the Artificial Neural Networks and Machine Learning—ICANN 2011, Espoo, Finland, 14–17 June 2011; Honkela, T., Duch, W., Girolami, M., Kaski, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 52–59.

33. Li, H.; Xu, Z.; Taylor, G.; Goldstein, T. Visualizing the Loss Landscape of Neural Nets. *arXiv* **2017**, arXiv:1712.09913.

34. Roy, H.; Chaudhury, S.; Yamasaki, T.; DeLatte, D.; Ohtake, M.; Hashimoto, T. Lunar surface image restoration using U-net based deep neural networks. *arXiv* **2019**, arXiv:1904.06683.

35. Lee, S.; Negishi, M.; Urakubo, H.; Kasai, H.; Ishii, S. Mu-net: Multi-scale U-net for two-photon microscopy image denoising and restoration. *Neural Netw.* **2020**, *125*, 92–103. [CrossRef] [PubMed]

36. Komatsu, R.; Gonsalves, T. Comparing U-Net Based Models for Denoising Color Images. *AI* **2020**, *1*, 465–486. [CrossRef]

37. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5967–5976. [CrossRef]

38. Jansson, A.; Humphrey, E.; Montecchio, N.; Bittner, R.; Kumar, A.; Weyde, T. Singing voice separation with deep U-Net convolutional networks. In Proceedings of the 18th International Society for Music Information Retrieval Conference, Suzhou, China, 23–27 October 2017.

39. Zhang, Z.; Liu, Q.; Wang, Y. Road Extraction by Deep Residual U-Net. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 749–753. [CrossRef]

40. Guo, S.; Yan, Z.; Zhang, K.; Zuo, W.; Zhang, L. Toward Convolutional Blind Denoising of Real Photographs. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 1712–1722.

41. Guo, B.; Song, K.; Dong, H.; Yan, Y.; Tu, Z.; Zhu, L. NERNet: Noise estimation and removal network for image denoising. *J. Vis. Commun. Image Represent.* **2020**, *71*, 102851. [CrossRef]

42. Tian, C.; Xu, Y.; Zuo, W. Image denoising using deep CNN with batch renormalization. *Neural Netw.* **2020**, *121*, 461–473. [CrossRef]

43. Zhang, K.; Zuo, W.; Zhang, L. FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. *IEEE Trans. Image Process.* **2018**, *27*, 4608–4622. [CrossRef]

44. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [CrossRef]

45. Buades, A.; Coll, B.; Morel, J.M. A Review of Image Denoising Algorithms, with a New One. *Multiscale Model. Simul.* **2005**, *4*, 490–530. [CrossRef]

46. Limshuebchuey, A.; Duangsoithong, R.; Saejia, M. Comparison of Image Denoising using Traditional Filter and Deep Learning Methods. In Proceedings of the 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phuket, Thailand, 24–27 June 2020; pp. 193–196. [CrossRef]

47. Fan, L.; Zhang, F.; Fan, H.; Zhang, C. Brief review of image denoising techniques. *Vis. Comput. Ind. Biomed. Art* **2019**, *2*, 7. [CrossRef]

48. Testolin, A.; Stoianov, I.; De Filippo De Grazia, M.; Zorzi, M. Deep Unsupervised Learning on a Desktop PC: A Primer for Cognitive Scientists. *Front. Psychol.* **2013**, *4*, 251. [CrossRef] [PubMed]

49. Lee, H.; Grosse, R.; Ranganath, R.; Ng, A. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; p. 77. [CrossRef]

50. Lan, R.; Zou, H.; Pang, C.; Zhong, Y.; Liu, Z.; Luo, X. Image denoising via deep residual convolutional neural networks. *Signal Image Video Process.* **2021**, *15*, 1–8. [CrossRef]

51. Seib, V.; Lange, B.; Wirtz, S. Mixing Real and Synthetic Data to Enhance Neural Network Training—A Review of Current Approaches. *arXiv* **2020**, arXiv:2007.08781.

52. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]

53. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2017**, arXiv:1412.6980.

54. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. In Proceedings of the IEEE International Conference on Computer Vision (ICCV 2015), Santiago, Chile, 7–13 December 2015; Volume 1502. [CrossRef]

55. Masters, D.; Luschi, C. Revisiting Small Batch Training for Deep Neural Networks. *arXiv* **2018**, arXiv:1804.07612.

56. Kandel, I.; Castelli, M. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express* **2020**, *6*, 312–315. [CrossRef]

57. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013.

58. *ASTM A1030/A1030M-11, Measuring Flatness Characteristics of Steel Sheet Products,Standard Practice for*; ASTM International: West Conshohocken, PA, USA, 2011.

59. Butterworth, S. On the theory of filter amplifiers. *Wirel. Eng.* **1930**, *7*, 536–541.

60. Oppenheim, A.; Schafer, R.; Buck, J.; Lee, L. *Discrete-Time Signal Processing*; Prentice Hall international editions; Prentice Hall: Hoboken, NJ, USA, 1999.

61. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [CrossRef]

62. Schafer, R.W. What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Process. Mag.* **2011**, *28*, 111–117. [CrossRef]

63. Weinberg, L.; Slepian, P. Takahasi's Results on Tchebycheff and Butterworth Ladder Networks. *IRE Trans. Circuit Theory* **1960**, *7*, 88–101. [CrossRef]

64. Daubechies, I. *Ten Lectures on Wavelets*; Society for Industrial and Applied Mathematics, 3600 University City Science Center: Philadelphia, PA, USA, 1992.

65. Meyer, Y. *Ondelettes et Opérateurs*; Number v. 1 in Actualités Mathématiques: Hermann, MS, USA, 1990.

66. Nielsen, M. On the Construction and Frequency Localization of Finite Orthogonal Quadrature Filters. *J. Approx. Theory* **2001**, *108*, 36–52. [CrossRef]

67. *ASTM A568/A568M-17a, Standard Specification for Steel, Sheet, Carbon, Structural, and High-Strength, Low-Alloy, Hot-Rolled and Cold-Rolled, General Requirements for*; ASTM International: West Conshohocken, PA, USA, 2017.

68. Peng, Z.; Wang, G. Study on Optimal Selection of Wavelet Vanishing Moments for ECG Denoising. *Sci. Rep.* **2017**, *7*, 4564. [CrossRef] [PubMed]

69. Galiano, G.; Velasco, J. On a nonlocal spectrogram for denoising one-dimensional signals. *Appl. Math. Comput.* **2014**, *244*, 859–869. [CrossRef]