

Letter

# High Sensitivity Snapshot Spectrometer Based on Deep Network Unmixing

Hui Xie <sup>†</sup>, Zhuang Zhao <sup>\*,†</sup>, Jing Han, Lianfa Bai and Yi Zhang

School of Electronic Engineering and Optoelectronic Technology, Nanjing University of Science and Technology, Nanjing 210094, China; xiehui@njjust.edu.cn (H.X.); eohj@njjust.edu.cn (J.H.); blf@njjust.edu.cn (L.B.); eo\_zhy441@njjust.edu.cn (Y.Z.)

\* Correspondence: Zhaozhuang@njjust.edu.cn

† These authors contributed equally to this work.

Received: 20 October 2020; Accepted: 3 December 2020; Published: 9 December 2020



**Abstract:** Spectral detection provides rich spectral–temporal information with wide applications. In our previous work, we proposed a dual-path sub-Hadamard-s snapshot Hadamard transform spectrometer (Sub-s HTS). In order to reduce the complexity of the system and improve its performance, we present a convolution neural network-based method to recover the light intensity distribution from the overlapped dispersive spectra, rather than adding an extra light path to capture it directly. In this paper, we construct a network-based single-path snapshot Hadamard transform spectrometer (net-based HTS). First, we designed a light intensity recovery neural network (LIRNet) with an unmixing module (UM) and an enhanced module (EM) to recover the light intensity from the dispersive image. Then, we used the reconstructed light intensity as the original light intensity to recover high signal-to-noise ratio spectra successfully. Compared with Sub-s HTS, the net-based HTS has a more compact structure and high sensitivity. A large number of simulations and experimental results have demonstrated that the proposed net-based HTS can obtain a better-reconstructed signal-to-noise ratio spectrum than the Sub-s HTS because of its higher light throughput.

**Keywords:** Hadamard transform spectrometer; snapshot HTS; neural network

## 1. Introduction

Spectral detection is widely used in industry, remote sensing, military and many other aspects. However, the traditional slit spectrometer has a contradiction between fast acquisition and high signal-to-noise ratio (SNR). If we want to improve the response speed of the system, it will inevitably reduce the acquisition time of the system, resulting in a decrease in the system SNR. In order to improve the SNR of detected spectra, researchers have introduced compressed sensing (CS) [1–3], computational slits [4,5] and deconvolution [6].

Compressed sensing is a representative method to realize snapshot acquisition. Compressed sensing was first proposed by Donoho, Candes and Tao in 2006 [1,2]. In the same year, Brady et al. of Duke University proposed a new spectral imaging technology called coded aperture snapshot spectral imager (CASSI) and developed several CASSI-based snapshot imaging spectrometers in the following years [7,8]. In the aspect of reconstruction algorithms, a greedy algorithm [4–6], L1 norm convex optimization algorithm [9–11], L1 norm non-convex optimization algorithm [12], Bayesian method [13–15] and other reconstruction algorithms have been developed. However, these reconstruction algorithms all have some drawbacks. Although the greedy algorithm has low computational complexity, the reconstruction effect is not ideal. The L1 norm minimization method has good reconstruction performance, but it has high computational complexity. The Bayesian method lacks a strict theoretical guarantee. Thus, we should develop new reconstruction algorithms.

Based on the idea of compressed sensing, researchers proposed a static multimode multichannel spectrometer (MMS), which takes advantage of the noise reduction performance of Hadamard coding and uses the non-negative least squares method to directly reconstruct the spectrum [16–18]. This method can achieve snapshot measurement and maintain the advantage of high light throughput. However, compressed sensing is an ill-posed inverse problem, and the measurement results are not stable enough, which limits its application.

Compared to multiplexing methods, these numerical methods belong to so-called ill-posed inverse problems and cannot obtain certain and robust SNR improvement. In our previous work [19], we proposed a dual-path sub-Hadamard-s snapshot Hadamard transform spectrometer (Sub-s HTS) based on the Hadamard transform spectrometer (HTS) [20,21]. In this Sub-s HTS, all incident light is encoded simultaneously by the encoding matrix, and an extra imaging path is designed to measure the intensity distribution of the scene. The spectral measurement problem is transformed into a positive definite problem so that a stable and reliable SNR enhancement can be obtained and maintain snapshot.

However, there are some problems in our previous work [19]. Firstly, the simultaneous acquisition of light intensity and dispersion image requires a dual optical path system, which will reduce the overall light throughput by half and decrease the SNR. Secondly, dual-camera images require pixel-level registration which will cause a lot of problems and affect the reconstruction results. Finally, the dual optical path system increases the overall system size and reduces reliability. Therefore, how to use a single camera to complete the reconstruction process while ensuring the quality of the reconstructed spectrum becomes a problem that must be solved.

In recent years, the convolution neural network has shown strong fitting ability in many fields. In order to improve the light throughput of the system and reduce its complexity, we designed a convolution neural network to recover the light intensity distribution through the overlapped dispersive spectra and realize the single-path high-sensitivity snapshot spectral measurement.

The main contributions of this paper are as follows:

- (1) A light intensity recovery network (LIRNet) to solve the problem of spectral image unmixing and realize the direct acquisition of light intensity distribution through an overlapped dispersive spectral image;
- (2) The feasibility of using reconstructed light intensity data for sub-Hadamard matrix spectral detection is proven;
- (3) Both simulated and experimental results demonstrated that the performance of the net-based scheme can achieve similar or even better reconstructed results compared with the dual-path scheme and can improve the compact of the system.

## 2. Design of Snapshot Spectrum Detection Framework

In our previous work [19], we proposed a dual-path snapshot Hadamard transform spectrometer. The implementation and scheme of the snapshot spectrometer are shown in Figure 1a,b, respectively.

The snapshot HTS contains two imaging paths: non-dispersive and dispersive imaging paths. The non-dispersive imaging path is employed to capture the light intensity at the coding aperture. The dispersive imaging path is employed to capture the overlapped dispersive spectra. We take the  $7 \times 7$  Hadamard matrix as an example to explain the spectral dispersion overlapping process in the snapshot spectrometer and show it in Figure 2.

In Figure 2,  $S_1$  is the 1st row of the Hadamard-S matrix,  $I_1$  is the 1st row light intensity distribution of the scene,  $f_1$  is the spectrum of the 1st row,  $f_{11}$  is the spectrum of the 1st pixel in the 1st row and  $m$  represents its range and the zeros in the spectrum represent shift invariant, which makes processing easy to understand. The measurement problems of the  $i$ th row of snapshot HTS are as follows:

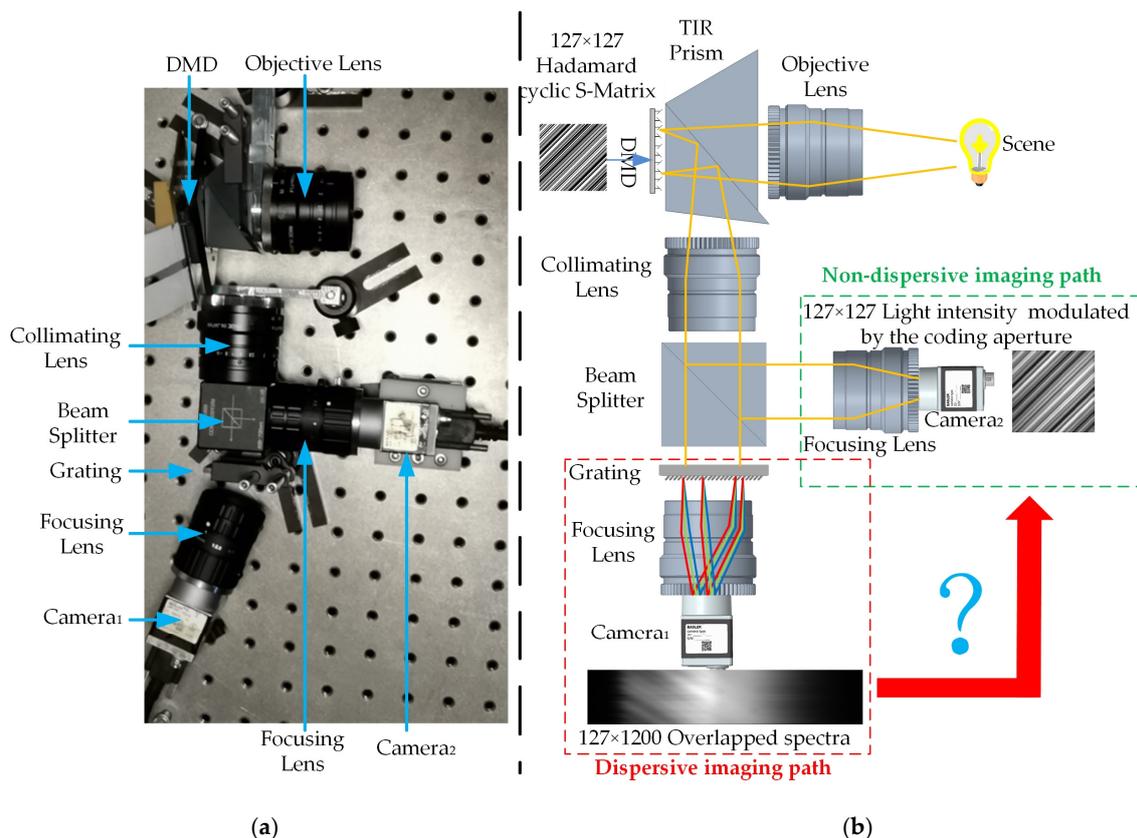
$$g_i = (S_i \circ I_i) f_i + n_{snap_i} \quad (1)$$

where  $g_i$  is the overlapping dispersion spectrum of the  $i$ th row,  $S_i$  is the  $i$ th row of Hadamard-S matrix,  $I_i$  is the  $i$ th row of light intensity distribution,  $f_i$  is the  $i$ th row of the spectrum to be measured and  $n_{snap_i}$  is the measurement noise in  $i$ th row measurement. If each row of the scene has the same spectrum, or we want to measure the average spectrum of each column, i.e.,  $f_1 = f_2 = \dots = f_n$  or  $f_{11} = f_{21} = \dots = f_{n1} = \frac{1}{n} \sum_{i=1}^n f_{i1}$ . Based on this assumption, the whole measurement of snapshot HTS can be simplified as:

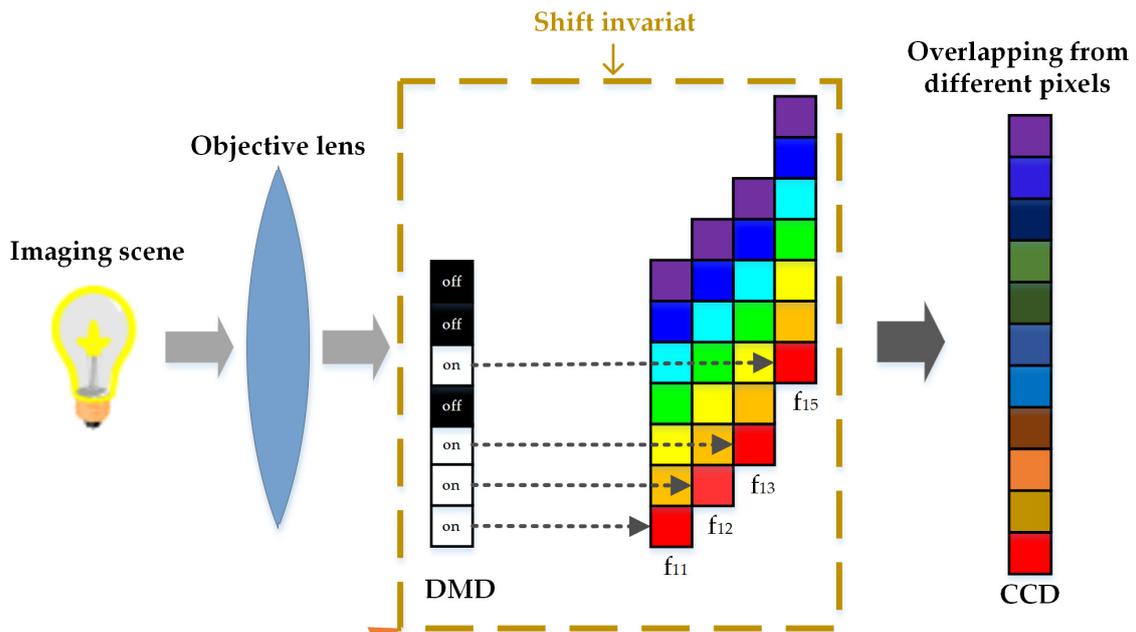
$$g = (S \circ I)f + n_{snap} = S_{snap}f + n_{snap} = (S - S_h)f + n_{snap} \quad (2)$$

where  $S_{snap}$  is the normalized modulation intensity distribution (sub-Hadamard-s matrix or Sub-matrix) and  $S_h$  is the difference between Hadamard-S matrix and normalized modulation intensity distribution. Thus, once the light intensity distribution  $S_{snap}$  is obtained, the spectra need to measure can be recovered through an inverse process, otherwise, we cannot recover the spectra.

As shown in Figure 2, the extra non-dispersive imaging path is added to measure the light intensity distribution. The extra imaging path increases the complexity of the system and weakens the intensity of the measured overlapped spectra. If we can recover the light intensity from the measured overlapped spectra such as the red arrow in Figure 1b, we can remove the extra imaging path and improve the performance of the system. We designed a light intensity recovery network (LIRNet) to recover the light intensity from the dispersive image. The overall structure of the network is shown in Figure 3.



**Figure 1.** The flowchart of dual-path snapshot Hadamard transform spectrometer (HTS). (a) Actual system diagram of HTS. (b) Principle system diagram of HTS.



Coding mask	Light intensity	Spectra of scene
$S = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$	$I = \begin{bmatrix} I_{11} & I_{12} & I_{13} & I_{14} & I_{15} & I_{16} & I_{17} \\ I_{21} & I_{22} & I_{23} & I_{24} & I_{25} & I_{26} & I_{27} \\ I_{31} & I_{32} & I_{33} & I_{34} & I_{35} & I_{36} & I_{37} \\ I_{41} & I_{42} & I_{43} & I_{44} & I_{45} & I_{46} & I_{47} \\ I_{51} & I_{52} & I_{53} & I_{54} & I_{55} & I_{56} & I_{57} \\ I_{61} & I_{62} & I_{63} & I_{64} & I_{65} & I_{66} & I_{67} \\ I_{71} & I_{72} & I_{73} & I_{74} & I_{75} & I_{76} & I_{77} \end{bmatrix}$	$f = \begin{bmatrix} f_{11} & f_{12} & f_{13} & f_{14} & f_{15} & f_{16} & f_{17} \\ f_{21} & f_{22} & f_{23} & f_{24} & f_{25} & f_{26} & f_{27} \\ f_{31} & f_{32} & f_{33} & f_{34} & f_{35} & f_{36} & f_{37} \\ f_{41} & f_{42} & f_{43} & f_{44} & f_{45} & f_{46} & f_{47} \\ f_{51} & f_{52} & f_{53} & f_{54} & f_{55} & f_{56} & f_{57} \\ f_{61} & f_{62} & f_{63} & f_{64} & f_{65} & f_{66} & f_{67} \\ f_{71} & f_{72} & f_{73} & f_{74} & f_{75} & f_{76} & f_{77} \end{bmatrix}$

$$\begin{aligned}
 & (S_{11} \circ I_{11}) \times [f_{11}^1 \ f_{11}^2 \ \dots \ f_{11}^m \ 0 \ 0 \ 0 \ 0 \ 0 \ 0] \\
 & + (S_{12} \circ I_{12}) \times [0 \ f_{12}^1 \ f_{12}^2 \ \dots \ f_{12}^m \ 0 \ 0 \ 0 \ 0 \ 0] \\
 & + (S_{13} \circ I_{13}) \times [0 \ 0 \ f_{13}^1 \ f_{13}^2 \ \dots \ f_{13}^m \ 0 \ 0 \ 0 \ 0] \\
 & + (S_{14} \circ I_{14}) \times [0 \ 0 \ 0 \ f_{14}^1 \ f_{14}^2 \ \dots \ f_{14}^m \ 0 \ 0 \ 0] \\
 & + (S_{15} \circ I_{15}) \times [0 \ 0 \ 0 \ 0 \ f_{15}^1 \ f_{15}^2 \ \dots \ f_{15}^m \ 0 \ 0] \\
 & + (S_{16} \circ I_{16}) \times [0 \ 0 \ 0 \ 0 \ 0 \ f_{16}^1 \ f_{16}^2 \ \dots \ f_{16}^m \ 0] \\
 & + (S_{17} \circ I_{17}) \times [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ f_{17}^1 \ f_{17}^2 \ \dots \ f_{17}^m]
 \end{aligned}$$

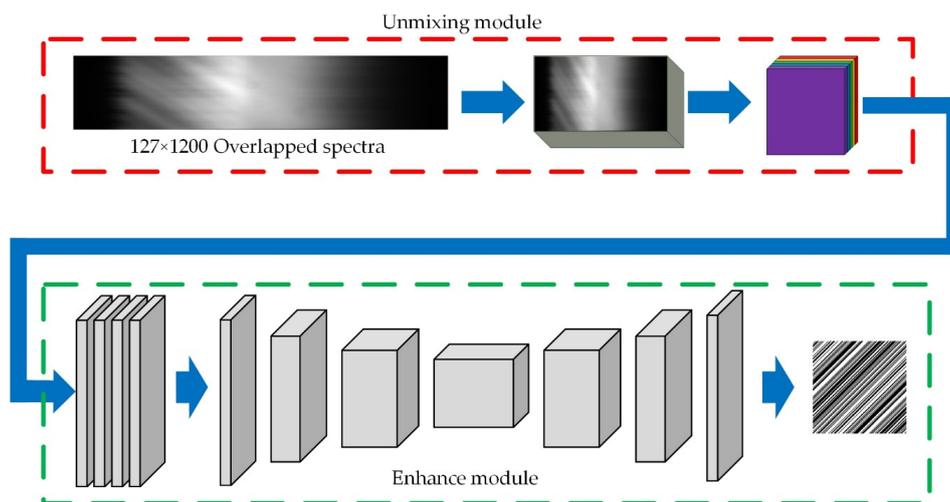
Measured overlapped spectra of 1<sup>st</sup> row  $g_1$

$$[I_{11} \circ f_{11}^1 \quad I_{11} \circ f_{11}^2 + I_{12} \circ f_{12}^1 \quad I_{11} \circ f_{11}^3 + I_{12} \circ f_{12}^2 + I_{13} \circ f_{13}^1 \quad L \quad I_{15} \circ f_{15}^m \quad 0 \quad 0]$$

$$g_1 = (S^1 \circ I_1) f_1 + \Delta n_1$$

$$\begin{aligned}
 f_1 &= [f_{11} \ f_{12} \ f_{13} \ f_{14} \ f_{15} \ f_{16} \ f_{17}] \\
 I_1 &= [I_{11} \ I_{12} \ I_{13} \ I_{14} \ I_{15} \ I_{16} \ I_{17}] \\
 S^1 &= [1 \ 1 \ 1 \ 0 \ 1 \ 0 \ 0]
 \end{aligned}$$

Figure 2. The coding processing of snapshot Hadamard transform spectrometer.



**Figure 3.** Structural sketch of the network model.

In LIRNet, we first designed an unmixing module (UM) to approximate the inverse process of convolution in the following Equation (3) which reduced the mixture of input pixels and output coarse unmixed spectral images. We then used an enhanced module (EM) to reconstruct the light intensity for sharper edges and better details.

### 3. Network Setup

As shown in Figure 3, in the raw spectral image, the intensity dispersion of each point is overlapped by grating, and the captured spectral image can be expressed as the overlapped image of each band spectral image. This process can be expressed by the convolution operator. Each pixel in the spectrum can be represented as:

$$f = \text{conv}(S, k) \quad (3)$$

where  $f$  is a single pixel in the spectral image,  $k$  is a one-dimensional convolution kernel in the spectral direction, the kernel size of convolution is the number of spectral bands and  $S$  is a high-dimensional spectral image.

From the perspective of spectral dimension, the light intensity of a point is the sum of its dispersive spectral intensity:

$$f_s = \sum_{i=1}^n S_i \quad (4)$$

where  $n$  is the number of spectral bands. This optical process is equal to the overlapping of a spectral image.

The UM contains five convolution layers and a deconvolution layer which extracts the primary features at the spectral dimension. The EM is an encoder–decoder network output feature map of UM, a further enhancement to supplement more visual information such as image contrast, brightness and texture features.

While training the network, we used two types of the loss function, intensity loss (IL) and spectrum loss (SL). IL adopts the mean square error (MSE) based on the online hard example mining (OHEM) [22] strategy as a loss function and SL adopts the SNR of the reconstructed spectrum as another loss function.

Euclidean distance is expressed as  $D_i = \|x_i - \hat{x}_i\|^2$ . The total number of samples is  $N$ , then:

$$\text{Loss}(x_i, \hat{x}_i) = \frac{1}{N} \sum_{i=1}^N D_i \quad (5)$$

where  $x_i, \hat{x}_i$  are real and reconstructed samples respectively. In this equation, samples with higher (high-frequency areas) and lower (low-frequency areas) errors are averaged for network training, which makes it difficult for the optimizer to optimize in the training stage. To make the optimizer focus on the harder samples, Wu et al. put forward a method of hard-sample mining for pixel-level classification tasks [22]. We adopted the idea and modified it to adapt our intensity approximation task. We selected pixel samples with higher loss values as hard negative samples for training and ignored others with lower loss values so that the network would not repeat learn samples in low-frequency areas and would improve the performance in high-frequency areas. In this case, the loss function can be expressed as:

$$Loss = \frac{1}{\sum_i^N 1\{D_i < t\}} \sum_i^N 1\{D_i < t\} D_i \quad (6)$$

where  $t$  is the threshold, which is the minimum loss value of the first half of the sample with a larger loss value.

In order to make the intensity approximation consistent with realistic physical processes better, we used spectrum loss to further fine-tune the network:

$$Loss_s = R - SNR(recon(\hat{x}), f) \quad (7)$$

where  $R$  is the reversal factor,  $recon$  is the reconstruction processing of overlapped spectrum and  $\hat{x}, f$  is the real spectrum.

We used 1650 groups of multispectral images with  $127 \times 127$  size for training, 350 groups for validation and 200 groups that were not involved in the training for testing. The images are padded to  $128 \times 128$  to train conveniently. Finally, a  $128 \times (128 + n)$  overlapped spectral image was constructed where  $n$  is the number of spectral bands. The light intensity images that were superimposed by all bands were used as the training label.

In the UM, the convolution layer with  $1 \times n$  convolution kernel was used for unmixing, and the convolution layer with  $3 \times 3$  convolution kernel and rectified linear unit (ReLU) [23] activation function was used to extract the primary features. In the enhancement module, the encoder–decoder network was used to extract the low-level and high-level semantic information of the unmixed feature map to enhance the unmixing image. The design of the enhance module structure is referred to as ERFNet [24], as shown in Table 1. The down-sampler layer is a pooling down-sampling process. The Non-bt-1d is the factorized convolution which decomposes the  $3 \times 3$  convolution into a pair of 1D convolutions, which constitutes a non-bottleneck 1D structure as shown in Figure 4. With a 33% reduction of parameters, we can achieve the same learning ability and accuracy as the traditional non-bottleneck structure, improve the network efficiency and contribute to real-time spectral analysis.

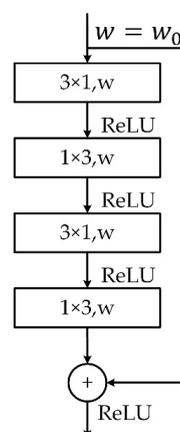


Figure 4. Non-bottleneck 1D structure.

**Table 1.** Network structure.

	Layer	Type	Out-F	Out-Res
SIMULATION	1	conv( $n \times 1$ )	1	$128 \times 128$
	2	deconv( $1 \times 1 \times n$ )	n	$128 \times 128$
	3	conv+ReLU	n/2	$128 \times 128$
	4	conv+ReLU	n/4	$128 \times 128$
	5	conv+ReLU	n/8	$128 \times 128$
	6	conv+ReLU	1	$128 \times 128$
ENCODER	7	Downsampler	16	$64 \times 64$
	8	Downsampler	64	$32 \times 32$
	9–13	$5 \times$ Non-bt-1D	64	$32 \times 32$
	14	Downsampler	128	$16 \times 16$
	15–22	$8 \times$ Non-bt-1D	128	$16 \times 16$
DECODER	23	deconv	64	$32 \times 32$
	24–25	$2 \times$ Non-bt-1D	64	$32 \times 32$
	26	deconv	16	$64 \times 64$
	27–28	$2 \times$ Non-bt-1D	16	$64 \times 64$
	29	deconv	C	$128 \times 128$

#### 4. SNR Analysis

In order to analyze the denoise performance of the net-based HTS, we set the sub-Hadamard matrix normalized intensity distribution as  $S_{snap}$ ;  $S_1$  is the difference between the network reconstructed light intensity and  $S_{snap}$  ( $S_1 \ll S_{snap}$ ). The reconstructed spectra can be written as follows:

$$\hat{f} = f + (S_{snap} - S_1)^{-1} n_s \quad (8)$$

Based on Equation (8), the SNR of the reconstructed spectrum can be expressed as:

$$SNR_{\hat{f}} = 10 \log \left( \frac{\frac{f^T f}{n_s^T (S_{snap}^T - S_1^T)^{-1} (S_{snap} - S_1)^{-1} n_s}}{\frac{f^T f}{n_s^T n_s}} \right) \quad (9)$$

According to our previous works,  $SNR_{\hat{f}}$  can be sampled as:

$$\begin{aligned} SNR_{\hat{f}} &\geq 10 \log \left( \frac{(n'_s)^T \left( (1-k)^2 S_{snap}^T S_{snap} + \frac{1-k}{k} k^2 S_{snap}^T S_{snap} \right) n'_s}{(n'_s)^T n'_s} \right) \\ &= 10 \log \left( \frac{(n'_s)^T S_{snap}^T S_{snap} n'_s}{(n'_s)^T n'_s} \right) + 10 \log(1-k) \end{aligned} \quad (10)$$

where  $n'_s = (S_{snap} - S_1)^{-1} n_s$ ,  $k$  denotes perturbations involving reconstruction errors and actual light intensity.

The results show that the reconstructed SNR of the network decreases by about  $10 \log\left(\frac{1}{1-k}\right)$  (dB) compared with the sub-Hadamard matrix which can accurately obtain the intensity distribution. Compared with a traditional slit spectrometer, it still has obvious advantages. In practice, the intensity of overlapped spectra in the single-path snapshot spectrometer is twice the intensity in the original dual-path snapshot spectrometer. Thus, the actual spectrum can be expressed as:

$$\hat{f} = 2f + (S_{snap} - S_1)^{-1} n_s \quad (11)$$

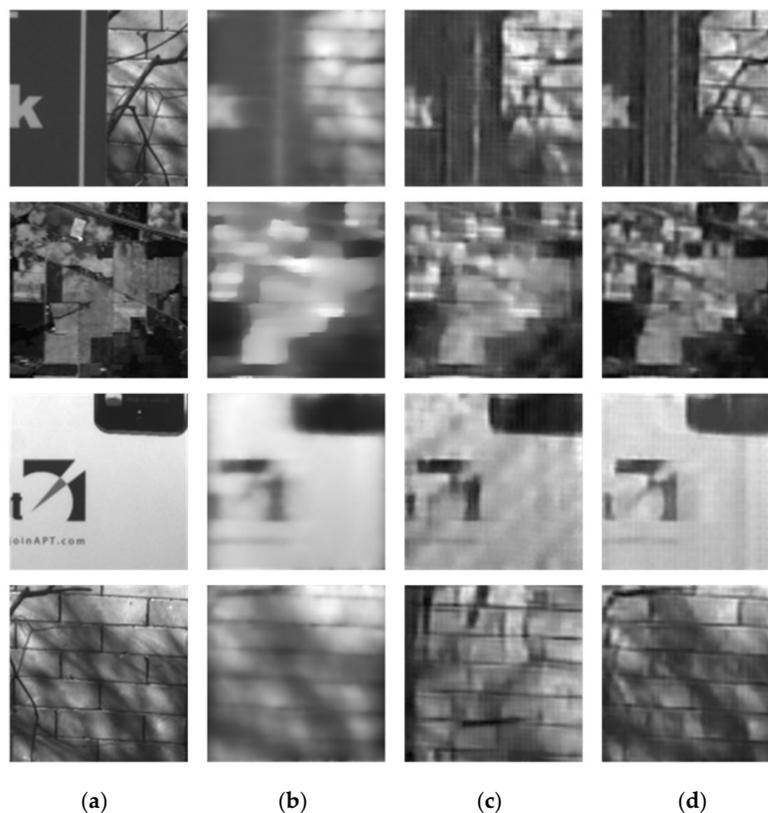
After a similar deduction, the final SNR can be expressed as:

$$SNR_{\hat{f}} \geq 10 \log \left( \frac{(n'_s)^T S_{snap}^T S_{snap} n'_s}{(n'_s)^T n'_s} \right) + 10 \log(1-k) + 10 \log(2) \quad (12)$$

Therefore, the proposed method has certain advantages over the dual-path Sub-s HTS.

## 5. Experiments and Simulations

In order to compare the reconstructed performance of different coding matrices, the full-1 matrix and Hadamard-S matrix are involved. Additionally, four frames CASSI [25] using two-step iterative soft threshold optimization (TwIST) are also involved. We present some reconstructed light intensity results that are different from the testing dataset [26,27] and calculate its peak signal-to-noise ratio (PSNR). The reconstructed results are shown in Figure 5. The resolution of reconstructed image is  $127 \times 127$ . The comparison of PSNR and reconstruction time of different scenes are shown in Table 2.



**Figure 5.** Reconstructed results of different methods. (The resolution of reconstructed image is  $127 \times 127$ ) (a) Ground truth. (b) Reconstructed results of coded aperture snapshot spectral imager (CASSI) 4 Frame. (c) Reconstructed results of the full-1 matrix (our method). (d) results of Hadamard-S matrix (our method).

**Table 2.** Comparison of peak signal-to-noise ratio (PSNR) and reconstruction time of different scenes.

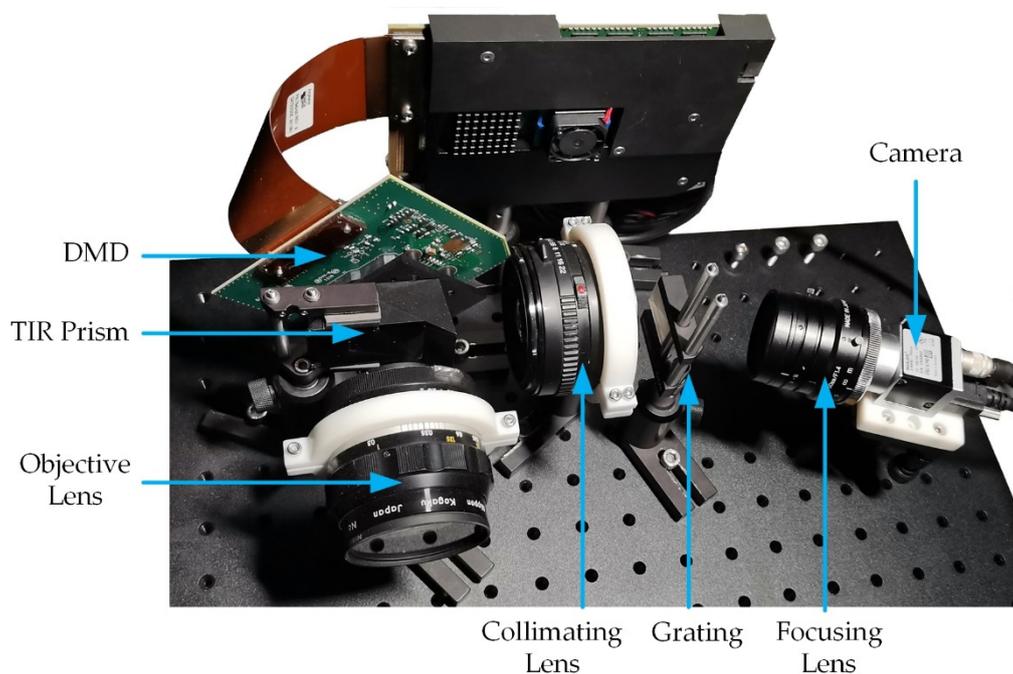
Image	CASSI 4 Frame	Full-1 (Ours)	Hadamard-S (Ours)
a	16.09 dB	22.37 dB	21.39 dB
b	12.39 dB	21.48 dB	21.13 dB
c	20.51 dB	20.82 dB	20.60 dB
d	19.44 dB	23.03 dB	22.90 dB
Reconstruction time	1.599 s	0.077 s	0.075 s

It can be seen that the performance of the proposed method is stable under different conditions, and the reconstruction quality of the proposed method is better than that of the four frames CS in the case of a single frame. The testing images are not involved in the training data, and the reconstructed results can also get the expected results, which shows that the network has good generalization ability. Additionally, we can find that the PSNR of reconstructed full-1 coding is better than Hadamard-S coding since the full-1 coding can use the full information and bring higher throughput. However, the full-1 matrix is an irreversible matrix and cannot ensure stable SNR boosting.

We bought a desktop on JD.com with Intel Core i7 8700K 4.3Ghz, manufactured by Intel, Santa Clara, CA, USA; 16GB memory, manufactured by Corsair Memory, Fremont, CA, USA; NVIDIA GeForce GTX 1060 6GB, manufactured by NVIDIA, Santa Clara, CA, USA; The reconstructed time is 0.075 s. It is 20 times faster than traditional methods without additional hardware acceleration. If we use a higher-performance GPU such as RTX TITAN, the reconstructed speed will be significantly improved, which can basically meet the real-time requirements.

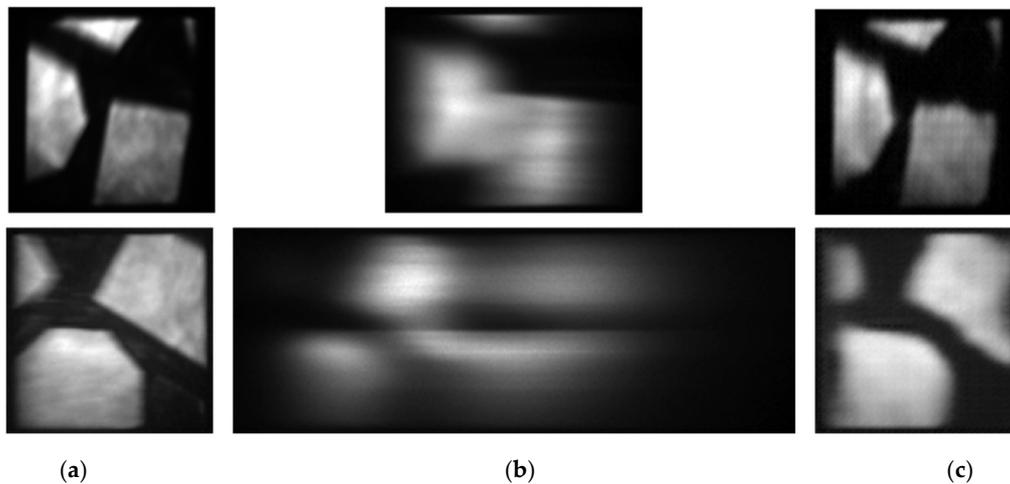
In the training strategy, we used the Adam algorithm for gradient optimization; the initial learning rate was 0.001, the whole training process lasted about 150 epochs when reaching convergence and batch size was set to 4. The convolution kernel size of the spectral dimension was modulated according to the dispersion length of the input spectral image.

To verify the effectiveness of reconstructed light intensity, we put the system into implementation and show it in Figure 6. The spectrum with band-pass filters and the whole spectrum were trained and tested in the actual optical system to test the influence of overlapping degrees on the reconstructed effect. In the implementation, the incident light is encoded by a digital micromirror device (DMD). The DMD (ED01N) has a resolution of  $1024 \times 768$  and a pixel size of  $13.68 \mu\text{m} \times 13.68 \mu\text{m}$ . The camera (Basler aca1920-155  $\mu\text{m}$ ) has a resolution of  $1920 \times 1200$  and a pixel size of  $5.86 \mu\text{m} \times 5.86 \mu\text{m}$ , and the dispersion grating is THORLABS GT25-03(300 Grooves/mm, 17.5 deg).



**Figure 6.** Implementation of single-path snapshot Hadamard transform spectrometer.

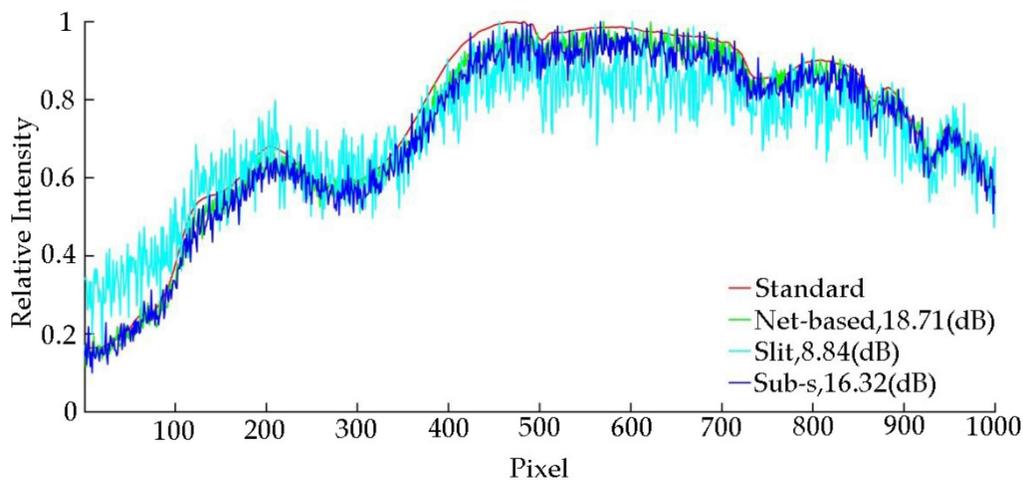
We also tested the light intensity reconstructed performance by increasing spectral bands, and the results are shown in Figure 7. A bandpass filter with 100 nm (400–500 nm) was used to truncate the spectrum, and the results are shown in the first row of Figure 7. The second row of Figure 7 is the result without any filter. The PSNR of reconstructed light intensity with filter is 25.76 dB and 20.26 dB without the filter.



**Figure 7.** Reconstructed results with different spectral bands. (a) Captured light intensity. (b) Measured overlapped dispersive spectra. (c) Reconstructed light intensity.

As shown in Figure 7, it can be seen that the proposed method can still reconstruct the light intensity well. However, with the increase of spectral bands, the overlapping becomes more serious, and the reconstructed quality will decrease.

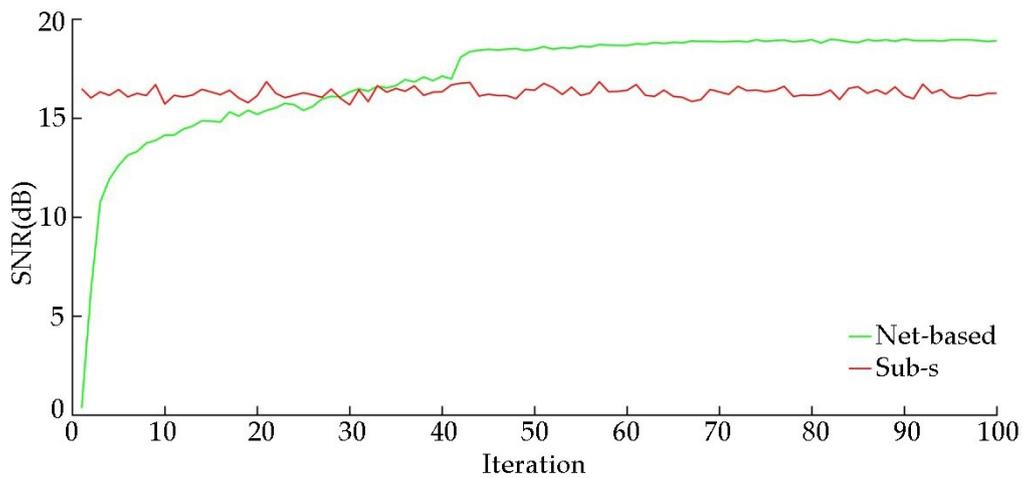
In practice, the Sub-s HTS will lose half throughput since it is the extra imaging path, which will have a negative impact on the SNR of the reconstructed spectrum. According to the theoretical deduction, the SNR decline is  $10\log(2)$  (dB). Thus, we compared the slit-based spectrometer, the Sub-s HTS [19] and the proposed net-based HTS through simulation and show the results in Figure 8.



**Figure 8.** Comparison of different methods.

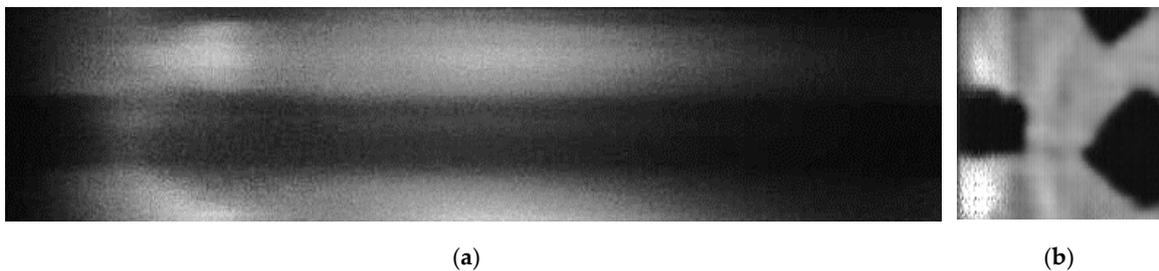
The results show that the proposed net-based HTS can achieve higher SNR results on the premise of reducing an optical path while maintaining snapshot. The SNR of dual-path Sub-s HTS is lower than the proposed net-based HTS which is in accord with the result of theoretical analysis.

The reconstructed quality of light intensity will affect the SNR of the reconstructed spectrum, and with the iteration of network, the quality of reconstructed light intensity will increase. Thus, we compared the SNR of reconstructed spectrum with the iteration of network and Sub-s HTS. The results are shown in Figure 9. It can be seen that as the iterations increase, the SNR of reconstructed spectra increases. As the quality of reconstruction increases, the advantage of light intensity gradually manifests, and the SNR is better than that of Sub-s HTS. This accords with the results of the previous theoretical derivation.

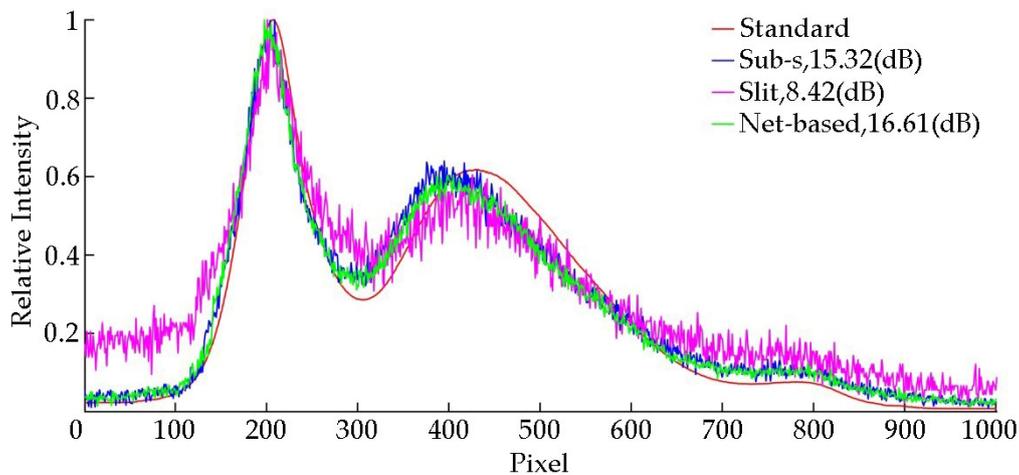


**Figure 9.** The SNR of the reconstructed spectrum with the iterations number of the network.

In order to evaluate the performance of the proposed net-based HTS, Sub-s HTS, slit-based spectrometer, we measured the SNR of each method with a low-level light source. In the experiment, we used an absorptive neutral density filter to cover an LED to simulate the low-level light source. We used the slit-based spectrometer to obtain the standard spectrum with the high light of LED and long exposure time. In order to compare the performance of all methods quantitatively, we set the same exposure time in all methods. The captured overlapped spectral image is shown in Figure 10, and the reconstructed spectral data are shown in Figure 11.



**Figure 10.** The results of actual experiments with 30 dB (analog gain of camera). (a) Overlapped dispersive spectra. (b) Reconstructed light intensity.



**Figure 11.** Experimental results of the proposed method, snapshot HTS and slit-based spectrometer with 100  $\mu$ s (exposure time of camera) and 30 dB (analog gain of camera).

The camera (Basler acA1920-155  $\mu\text{m}$ ) provides a maximum analog gain setting of 36 dB that increases the brightness of the images output. However, with the increase of analog gain, a lot of electronic noise and temporal dark noise will be introduced. It will be a great influence on the final image quality.

In Figure 11, it can be seen that due to the advantage of light throughput, the net-based HTS has a better SNR than the Sub-s HTS in the case of low-level light environment and strong noise interference. This proves that our system has stronger anti-noise capability. In the case of limited detector capability, our system can greatly improve the SNR of the detected spectrum.

## 6. Conclusions

In this paper, we propose a light intensity recover network (LIRNet) to obtain the original image from overlapped spectra and construct a network-based single-path snapshot Hadamard transform spectrometer (net-based HTS) system. Its advantages over our previous work (Sub-s HTS) lie in the following three aspects. First, we proposed a neural network LIRNet to recover the light intensity distribution from the overlapped dispersive spectra, where the calculation speed is 20 times faster than the traditional methods. Second, the proposed net-based HTS can solve the drawbacks of Sub-s HTS such as the pixel-level registration and the loss of light throughput. Third, thanks to the simple structure of the single-path optical system and the advantage of light throughput, the net-based HTS can obtain a better performance than Sub-s HTS, which further improves its practicability. Additionally, our system can greatly improve the SNR of detected spectrum with limited detector performance.

At the same time, the proposed idea of spectral unmixing can be further developed in the field of computational spectral imaging. We believe that the proposed framework can promote the development and applications of spectral imaging, with reduced hardware and software complexity. In our next research, we will introduce it into hyperspectral imaging to obtain clearer hyperspectral images, which will be an interesting avenue for future work.

**Author Contributions:** Methodology, Validation, Writing—original draft, Investigation, Data curation, H.X.; Methodology, Writing—review, Investigation, Data curation, Z.Z.; Conceptualization, Writing—review & editing, Project administration, Funding acquisition, J.H.; Supervision, Funding acquisition, L.B.; Writing—review & editing, Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant Nos. 61727802, 61901220, 2019K216).

**Acknowledgments:** We thank Xu Wang, Haocun Qi, Jiang Yue, Xiaoyu Chen and Enlai Guo for technical support.

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

## References

1. Candes, E.J.; Romberg, J.; Tao, T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory* **2006**, *52*, 489–509. [[CrossRef](#)]
2. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306. [[CrossRef](#)]
3. Brady, D.J.; Gehm, M.E. Compressive imaging spectrometers using coded apertures. *Vis. Inf. Process.* **2006**, *6246*, 62460A-1–62460A-9.
4. Huang, H.; Makur, A. Backtracking-based matching pursuit method for sparse signal reconstruction. *IEEE Signal Process. Lett.* **2011**, *18*, 391–394. [[CrossRef](#)]
5. Donoho, D.; Tsaig, Y.; Drori, I.; Starck, J.L. *Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit*; Technical Report; Department of Statistics, Stanford University: Stanford, CA, USA, 2006.
6. Needell, D.; Tropp, J. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* **2009**, *26*, 301–321. [[CrossRef](#)]
7. Wagadarikar, A.; John, R.; Willett, R.; Brady, D.J. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.* **2008**, *47*, B44–B51. [[CrossRef](#)]

8. Cao, X.; Yue, T.; Lin, X.; Lin, S.; Yuan, X.; Dai, Q.H.; Carin, L.; Brady, D.J. Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world. *IEEE Signal Process. Mag.* **2016**, *33*, 95–108. [CrossRef]
9. Gill, P.; Wang, A.; Molnar, A. The in-crowd algorithm for fast basis pursuit denoising. *IEEE Trans. Signal Process.* **2011**, *59*, 4595–4605. [CrossRef]
10. Daubechies, I.; Defrise, M.; De-Mol, C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **2004**, *57*, 1413–1457. [CrossRef]
11. Bioucas-Dias, J.M.; Figueiredo, M.A.T. A new TwIST: Two-step iterative shrinking/thresholding algorithms for image restoration. *IEEE Trans. Image Process.* **2007**, *16*, 2992–3004. [CrossRef] [PubMed]
12. Chartrand, R.; Yin, W. Iteratively reweighted algorithms for compressive sensing. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, NV, USA, 31 March–4 April 2008; pp. 3869–3872.
13. Tipping, M.E. Sparse bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244.
14. Tipping, M.E.; Faul, A.C. Fast marginal likelihood maximization for sparse bayesian models. In Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics, Key West, FL, USA, 1 January 2003; pp. 3–6.
15. Schniter, P.; Potter, L.C.; Ziniel, J. Fast bayesian matching pursuit. In Proceedings of the Workshop on Information Theory and Applications, La Jolla, CA, USA, 1 January 2008; pp. 326–333.
16. Wagadarikar, A.; Gehm, M.E.; Brady, D.J. Performance comparison of aperture codes for multimodal, multiplex spectroscopy. *Appl. Opt.* **2007**, *46*, 4932–4942. [CrossRef] [PubMed]
17. Gehm, M.E.; McCain, S.T.; Pitsianis, N.P.; Brady, D.J.; Potluri, P.; Sullivan, M.E. Static two-dimensional aperture coding for multimodal, multiplex spectroscopy. *Appl. Opt.* **2006**, *45*, 2965. [CrossRef] [PubMed]
18. Fernandez, C.; Guenther, B.D.; Gehm, M.E.; Brady, D.J.; Sullivan, M.E. Longwave infrared (LWIR) coded aperture dispersive spectrometer. *Opt. Express* **2007**, *15*, 5742–5753. [CrossRef]
19. Zhao, Z.; Bai, L.; Han, J.; Yue, J. High-SNR snapshot multiplex spectrometer with sub-Hadamard-S matrix coding. *Opt. Commun.* **2019**, *453*, 124322–124328. [CrossRef]
20. Yue, J.; Han, J.; Zhang, Y.; Bai, L. Denoising analysis of hadamard transform spectrometry. *Opt. Lett.* **2014**, *39*, 3744–3747. [CrossRef]
21. Yue, J.; Han, J.; Li, L.; Bai, L. Denoising analysis of spatial pixel multiplex coded spectrometer with hadamard, H.-matrix. *Opt. Commun.* **2018**, *407*, 355–360. [CrossRef]
22. Wu, Z.; Shen, C.; van den Hengel, A. High-performance semantic segmentation using very deep fully convolutional networks. *arXiv* **2016**, arXiv:1604.04339.
23. Xu, B.; Wang, N.; Chen, T.; Li, M. Empirical evaluation of rectified activations in convolutional network. *arXiv* **2015**, arXiv:1505.00853.
24. Romera, E.; Alvarez, J.M.; Bergasa, L.M.; Arroyo, R. ERFNet: Efficient residual factorized ConvNet for real-time semantic segmentation. *IEEE Trans. Intell. Transp. Syst.* **2018**, *19*, 263–272. [CrossRef]
25. Kittle, D.; Choi, K.; Wagadarikar, A.; Brady, D.J. Multiframe image estimation for coded aperture snapshot spectral imagers. *Appl. Opt.* **2010**, *49*, 6824–6833. [CrossRef] [PubMed]
26. Indian Pines. Available online: <http://lesun.weebly.com/hyperspectral-data-set.html> (accessed on 5 December 2020).
27. Chakrabarti, A.; Zickler, T. Statistics of real-world hyperspectral images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Colorado Springs, CO, USA, 23 June 2011; pp. 193–200.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).