Tamminen, Aleksi; Palli, Samu Ville; Ala-Laurinaho, Juha; Taylor, Zachary D.

Millimeter- and Submillimeter-Wave Imaging Through Dispersive Hologram and Deep Neural Networks

Published in:
IEEE Transactions on Microwave Theory and Techniques

DOI:
10.1109/TMTT.2022.3166176

Published: 01/06/2022

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Tamminen, A., Palli, S. V., Ala-Laurinaho, J., & Taylor, Z. D. (2022). Millimeter- and Submillimeter-Wave Imaging Through Dispersive Hologram and Deep Neural Networks. *IEEE Transactions on Microwave Theory and Techniques*, *70*(6), 3281-3290. https://doi.org/10.1109/TMTT.2022.3166176

TMTT-2021-11-1467

Millimeter- and submillimeter-wave imaging through dispersive hologram and deep neural networks

Aleksi Tamminen, *Member, IEEE*, Samu-Ville Pälli, *Student Member, IEEE*,
Juha Ala-Laurinaho, and Zachary D. Taylor, *Member, IEEE*.

***Abstract*— We present imaging results with dual-band millimeter- and submillimeter-wave hologram and deep neural networks. The imaging method uses a single transceiver, which interrogates the region of interest through a dispersive transmission-type hologram. The hologram was designed to cover two bands 50-75 GHz and 220-330 GHz. Two separate single-transceiver imaging experiments were carried out with two test objects translated in the region of interest at 101 × 101 locations. Neural networks were trained to images of the test objects with wide-band reflection spectra from the region of interest as the input. The deep neural networks were based on deconvolutional layers that mapped the latent information of the test objects in the spectra to image pixel values. The two ~10-cm test objects were imaged in 200 × 200 mm^2 and 300 × 300 mm^2 field-of-view at 600 mm from the hologram aperture (19° to 28° angular field-of-view). The experimental resolution was estimated from point-spread functions extracted from the predicted images. The full width at half maximum resolution was 21 mm and 16.5 mm, for the 50-75 GHz and 220-330 GHz bands, respectively. These are close to the theoretical limits of 25 to 19 mm, for the lower band, and 19 to 16 mm for the higher band as predicted with hologram aperture size and edge taper. Augmented reflections were constructed from corner-cube measurements to evaluate the ability to predict images of vast collection of objects. The results with augmented data show performance comparable with the experimental ones with limited test object space. The latent representations for both the experimental and augmented data indicate sparsity – a demonstration of feasibility to generalize from reflection spectra to images. The performance of the developed imaging technique is in par with the current, multichannel state of art and has the advantage of substantially reduced hardware complexity.**

***Index Terms*— Imaging, hologram, deep neural network, submillimeter wave**

I. Introduction

THE partial transparency of common materials and high sensitivity to water content are some of the key parameters that continue to drive interest and development in millimeter- and submillimeter-wave imaging technology. In addition to numerous scientific applications, imaging is used in personnel screening, medical diagnosis, and non-destructive testing. Many applications under research require real-time imagery, which often includes increasing sensor count or expanding optomechanics complexity. Despite significant advances in transceiver development, the unit cost for a sensor at these bandwidths remains high.

The image-forming technologies used in submillimeter-wave real-time screening systems are often based on electronic or mechanical beam scanning combined with a quasioptical coupling of energy from object to sensors. In scanned systems, the object plane focal point is steered by mirrors actuated by mechanical drives. A broad range of imaging systems based on passive radiometry and active imaging radars have been demonstrated and are commercially available [1].

A challenge in real-time imaging systems for personnel screening comes from the required signal-to-noise ratio (SNR), which is proportional to the integration time of the sensor. The sensors are multiplexed over the field-of-view to form the full image [2], and the sensor count is often maximized to achieve fast and high-SNR systems. Passive submillimeter-wave imaging radiometric sensor arrays utilizing dense, high device count have been developed both for room-temperature and cryogenic operation. For example, imaging with arrays of >1000 kinetic-inductance bolometers have been demonstrated with sufficient spatial resolution and field-of-view without moving optomechanics [3]. However, cooled systems require elaborate cryogenic technology that increases system cost and footprint as well as introduces startup/initialization time that may be up to 24-hours long.

Active imaging radars have been developed around linear arrays with scanning systems that sweep the 1D array over a 2D region of interest (RoI) [4]. Active imaging systems may apply sensor arrays with controllable phase shifts resulting in electronic beam steering that avoids actuating large physical elements. Phase shifts can also be applied for post detection in a multi-sensor system for similar effect [5]. Imaging systems with electronic beam steering may include thousands of sensors operating in a coherent way [6]. Although the cost of transceiver technology at millimeter waves has come down thanks to mass production for telecommunications and automotive radar, the technology remains prohibitively expensive, especially at higher millimeter wave frequencies and at submillimeter waves.

"Single-pixel camera" refers to technology where an imaging system relies on a single transceiver that is then multiplexed to different modes to interrogate the target in specific ways. Transmittance images of 2D test objects have been reported [7]-

Manuscript received November 25, 2021; revised February 21, 2022; accepted March 13, 2022. This work has been funded as research project ADENN through Academy of Finland programme ICT 2023 Beyond 5G (decision number 319770). *Corresponding author: A. Tamminen, phone: +358505137989, email: aleksi.tamminen@aalto.fi.*

A. Tamminen, S. Ville-Pälli, J. Ala-Laurinaho, and Z. D. Taylor are with Aalto University Department of Electronics and Nanoengineering, MilliLab, Espoo, Finland.

[10]. The modes can be created with masks that introduce spatial modulation into the illumination on the target. In optical wavelengths, the multiple masks can be created in fast succession by using spatial light modulators such as microelectromechanical mirror arrays or liquid-crystal devices. Our work presented here is similar to single-pixel camera as it bears only one transceiver and a quasioptical element with analogous function to the masks used in single-pixel experiments. There is no need for active control as the dispersive hologram passively varies the target illumination over a frequency sweep.

Several emerging technologies have been suggested to limit the number of transceivers while keeping the mechanical beamsteering at minimum. Such technologies are millimeter-wave holographic imaging methods [11], [12], frequency-diverse images based on highly over-moded millimeter-wave resonating antenna elements [13], and MIMO imaging systems [14], and CMOS-based active metamaterials [15]. At millimeter and submillimeter waves, the spatial light modulators have been demonstrated, such as the photo-injected Fresnel-zone plate antenna [16], [17]. In these imaging methods, beam steering is carried out as a single-mode process, i.e., a single diffraction-limited spot is scanned across the target and multiplexed in time as the scan proceeds. Imaging based on quasi-random illumination of the region of interest has been demonstrated at millimeter waves in [13]. Frequency-diverse millimeter-wave imaging with a telecommunications transceiver array is demonstrated in [18]. The designs are rather complex and typically also an accurate model of the radio-wave propagation within the system is needed for the inversing task in image formation. Our work takes advantage of the frequency diversity of the dispersive hologram in creating complex radiation patterns and it also uses deep neural network to carry out the inversing task from one-dimensional spectra to images. This can be considered a great advantage compared to the more complex imaging methods.

At optical wavelengths, holographic imaging process has been carried out without the knowledge of the wave propagation in the optical arrangement. A neural network (NN) demonstrated its ability to recover amplitude and phase from holograms in microscopic application [19]. Also, a deep learning method was proposed to predict a phase profile from diffracted field intensity [20]. Computational imaging and image segmentation using both convolutional [21] and deconvolutional [22], [23] neural networks are applied with sparse or latent representations of images as the input. Deconvolutional neural network is selected as the neural-network topology in our work due to its property to expand latent representations into images.

At millimeter waves, machine learning has been used in non-destructive testing [24] and in biomedical sample classification [25]. As tangential work to that presented here, neural networks have been utilized to design diffractive THz components [26]-[28]. The NN was used to design output spectra from metamaterial structures [29]. The THz wave propagation is modeled with neural networks in a super-resolution application of a scanned focused THz imaging system [30]. Millimeter-wave imaging of simple reflective targets has been presented in [31], where machine learning was applied to reconstruct the image. The abovementioned machine-learning applications are used to aid in designing components or as an additional layer of image classification, whereas our novel method is an end-to-end solution to computationally form images from reflection spectra.

We present the first results of millimeter- and submillimeter-wave imaging based on dispersive hologram structure and deconvolutional neural networks. The results are obtained at two separate bands: 50-75 GHz and 220-330 GHz. The dispersive hologram distributes the waves on the target as the transmitter frequency is swept. The back reflected waves are received as spectra that are unique to the location and reflectivity distribution of the test objects. The neural networks are based on deconvolutional layers, and they have been earlier used in semantic segmentation task to expand a latent representation from a scene into an image [23]. Our key hypothesis is that the reflection spectrum from the region of interest through the dispersive hologram is a latent representation that conveys the image information. The new imaging method allows for imaging without mechanical scanning and can potentially reduce the sensor count and optomechanical complexity in imaging systems. Our experiments at millimeter- and submillimeter waves show that the cross-range resolution in the imaging task are in line with the diffraction limit. Earlier, we have demonstrated localization based on similar dispersive element combined with terahertz time-domain spectrometer operating at 0.1-2 THz [32].

In the following, we will discuss the imaging method and the dispersive hologram in Section II, we present the data from measurements with extended targets as well as with augmented data composed of measured reflections from a corner-cube reflector in Section III, the deconvolutional neural network is introduced in Section IV, and the imaging experiment results are shown in Section V. Finally, we end up with conclusions.

II. The Imaging Method

In the proposed imaging method, the illumination is directed towards the region of interest (RoI) through a transmission-type hologram. With a proper design, the modulated field is both directive and dispersive; there is minimal spatial (transverse correlation) between spectral components and the majority of energy within the operational band is directed towards the region of interest. The RoI is thus illuminated with a high amplitude, frequency-diverse, spatially varying field. In our case, the imaging system consists of a single transceiver; a millimeter-wave extender coupled to a vector-network analyzer (VNA), which acquires back reflection from the RoI through the hologram. The reflection includes the reflectivity map of the object weighted with the spatial field distribution, which is highly varying as function of frequency. Fig. 1 shows the arrangement of the hologram and the field dispersion in the RoI.

The proposed method is based on the following hypothesis: If the illuminating field varies enough spatially across the bandwidth, the measured wide-band reflection will carry enough information to form an image of the object. The image formation is performed through a trained deep neural network. A deconvolutional neural network was selected for the image formation tasks, as it can be used to construct the images based on latent frequency-domain representation of the object [33].

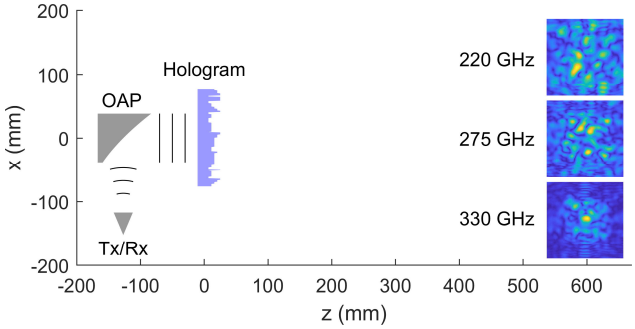

Fig. 1. Frequency-diverse illumination of the region of interest through a hologram. The field in the region of interest is changing rapidly as function of frequency.

Two setups operating at two different bands were constructed: WR-15 at 50-75 GHz and WR-3.4 at 220-330 GHz. The imaging system was fed by Pickett-Potter horn antennas and radiation was collimated by a 76.2-mm offset-axis parabolic mirror (OAP). The OAP mirror's 127-mm effective focal length resulted in approximately radially symmetric amplitude in the collimated beam with an edge taper of 6 dB at WR-15 and 28 dB at WR-3.4. These edge tapers are approximated with Gaussian beams. Following the Gaussian-beam theory, at 600-mm distance, the edge tapers correspond image resolutions of 25 to 19 mm and 19 to 16 mm, for WR-15 and WR-3.4, respectively [34].

A. The Transmission-Type Dielectric Hologram

The design process of the hologram is presented in detail in [35] and it is analogous to zone-plate lens design described, e.g., in [36]. During the hologram synthesis process, a goal field was defined first at a specific frequency, f_{design}, in the region of interest ($z = z_{\mathrm{RoI}}$):

$$E_{\mathrm{goal}}(x_i, y_j, f_{\mathrm{design}}, z_{\mathrm{RoI}}) = \sqrt{R_{\mathrm{i,j}}}\,\mathrm{e}^{j2\pi P_{\mathrm{i,j}}}, \tag{1}$$

where random variables are used as the amplitude $\sqrt{R_{\mathrm{i,j}}}, R_{\mathrm{i,j}} \in U(0,1)$ and phase $P_{\mathrm{i,j}} \in U(0,1)$. The hologram transforms a collimated beam, E_{c}, into the goal field, E_{goal}, so that the field at the target is quasirandom and changing as function of frequency. In the design of the hologram, the collimated beam incident to the hologram is assumed Gaussian: $E_{\mathrm{c}}(r) = \mathrm{e}^{-r^2/{w_0}^2}$ where r and w_0 is the radial coordinate and frequency dependent beam waist respectively. A physical-optics method was used to calculate the propagation of the random goal field to the 76.2 mm × 76.2 mm hologram surface located at z = 0 mm. As a result, the diffracted goal field truncated by the hologram aperture is obtained ($E_{\mathrm{goal}}(f_{\mathrm{design}}, z_{\mathrm{a}})$). The design process of the hologram structure is shown in Fig. 2. Transmission function, T, through the hologram is realized so that the collimated beam is transformed into approximation of the field when it propagates through the hologram. The transmission function is

$$T(f_{\mathrm{design}}) = \frac{E_{\mathrm{goal}}(f_{\mathrm{design}}, z_{\mathrm{a}})}{E_{\mathrm{c}}}. \tag{2}$$

In our case, the transmission function of the hologram was chosen to operate on the electric field phase:

$$T_{\mathrm{h}}(f_{\mathrm{design}}) = \mathrm{e}^{j\Gamma(\angle T(f_{\mathrm{design}}))}, \tag{3}$$

where $\Gamma(\cdot)$ is a quantization function. The design frequency was $f_{\mathrm{design}} = 120$ GHz. The selected transmission function does not consider amplitude, which is affected only by dielectric loss and reflections within the structure. The quantization function controls relief height. Larger reliefs are easier to manufacture, with relaxed mechanical tolerances. An optimal quantization avoids high-aspect-ratio details that may compromise realization of the desired transmission. An optimal quantization function is essential for achieving the desired dispersive property. In our case the quantization function is

$$\Gamma(\angle T) = (m+q)\mathrm{round}(\angle T/q), \tag{4}$$

where $q = \pi/2$ is the quantization step and $m = 1.9\pi$ is a phase-shift offset. The phase-shift offset increases the thickness of the relief, which allows relative phase differences to vary strongly as a function of frequency across 50-75 GHz and 220-330 GHz. The choice of the design frequency, quantization step, and phase-shift offset are discussed in [35], [37]. Fig. 2 shows the design process to determine the transmission function.

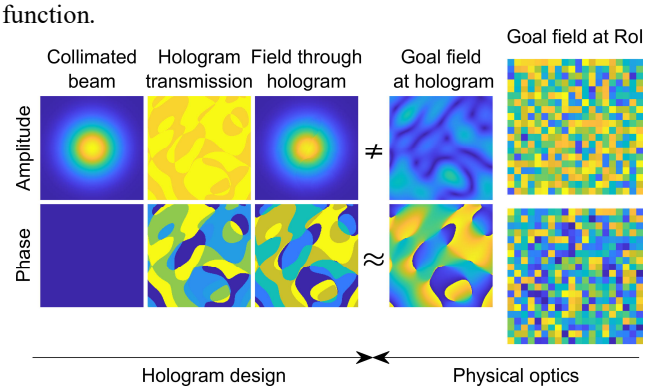

Fig. 2. Design process of the hologram. On left is the collimated field from the OAP mirror and on right is the random goal field at the RoI. The goal field phase at the hologram is approximated by designing the hologram transmission accordingly. The hologram affects the transmitted amplitude only little, and it remains almost Gaussian after transmission.

In this work, the hologram was a dielectric slab with a surface relief and negligible material dispersion. The height of the surface relief was modulated so that the desired phase shift was introduced at the design frequency. The transmission through the hologram is assumed locally to follow the 1D dielectric-slab model

$$T = \frac{\tau_1 \tau_2 e^{-j\beta h}}{1 - \rho_1 \rho_2 e^{-2j\beta h}}, \tag{5}$$

where τ_1, τ_2 are transmission and ρ_1, ρ_2 are reflection coefficients through the front and back surfaces of the hologram, β is the propagation constant in the dielectric and h is the height of the hologram surface relief. In addition to the phase modulation, (4) results in small amplitude modulation as well (see Fig. 2, "Hologram transmission"). The hologram design process only considers the phase shift in transmission from the planar wavefront in the collimated beam to the approximate of the phase in the goal field. The hologram was

manufactured of Rexolite 1422, which has relative dielectric permittivity of $\varepsilon_r \approx 2.52\text{-}j0.0005$ [38]. Considering the permittivity, hologram surface relief height is given as

$$h = \frac{-\Gamma(\angle T)}{(\sqrt{\varepsilon_r} - 1)k_0}, \tag{6}$$

where ε_r is the relative permittivity of Rexolite and k_0 is the free-space wavenumber at the design frequency.

The simple design principles presented here do not account for the more complex propagation effects in the hologram, such as edge effects near height change in the hologram. The hologram was manufactured with CNC milling, and some of the smaller synthesized relief features, described by (6) could not be accurately reproduced due to the 1-mm mill-bit diameter. The manufactured surface profile was calculated by machining simulator, which enforces constant cutter load and approximates the given relief with boundary conditions set by the available cutter-piece dimensions. The differences between designed and realized relief patterns appeared to have limited effects on the desired goal fields and performance. An isometric view of the 3D hologram design is shown in Fig. 3.

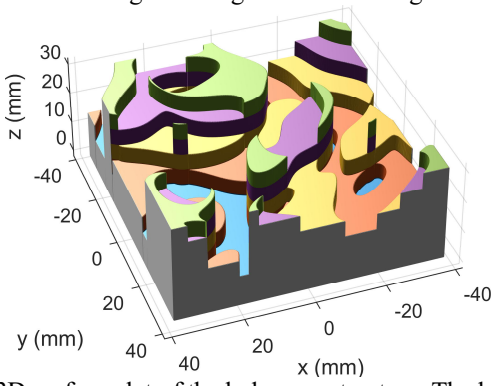

Fig. 3. 3D surface plot of the hologram structure. The hologram is discretized in five different height levels (shown in colors).

From (4) and (6), the hologram has five height levels in the surface relief at the design frequency, which correspond to phase shifts of $0,1,2,3,4\times(\pi/2+1.9\pi)$. The 76.2 mm × 76.2 mm hologram was milled into a 30-mm thick slab of Rexolite. The overall dimension of the slab was 94 mm × 94 mm to leave room for mounting screws. The expected and measured phase profiles have been studied with near-field imaging at 275 GHz, and the fabricated hologram delivers excellent match with average error of less than 6° [37].

B. Imaging Setup

The 220-330-GHz imaging setup is shown in Fig. 4 The setup consists of a submillimeter-wave extension VNAX WR-3.4 (Virginia Diodes Inc.) connected to network analyzer N5225B PNA (Keysight Technologies). A Pickett-Potter horn (Radiometer Physics GmbH) was used as the feed for an offset parabolic mirror (Edmund Optics) collimating the beam. The hologram produced a frequency-dependent beam pattern in RoI at 600 mm from the hologram surface. The field in the region of interest was characterized with near-field measurements and reported in [35]. In the imaging experiment presented here, the two Aalto University logos made of porous copper foam were mounted on the near-field scanner in the region of interest. The copper foam was 5-mm thick, and its pore size was 1-2 mm. The pore size is of the order of the wavelength, which resulted in strong isotropic scattering of the incident field. The 135 mm × 100 mm and 154 mm × 100 mm logos have smallest detail of about 6 mm. The measured reflection from the setup was used as an input and binary, affine-transformed photographs of the logos were used as the ground truth in neural-network training.

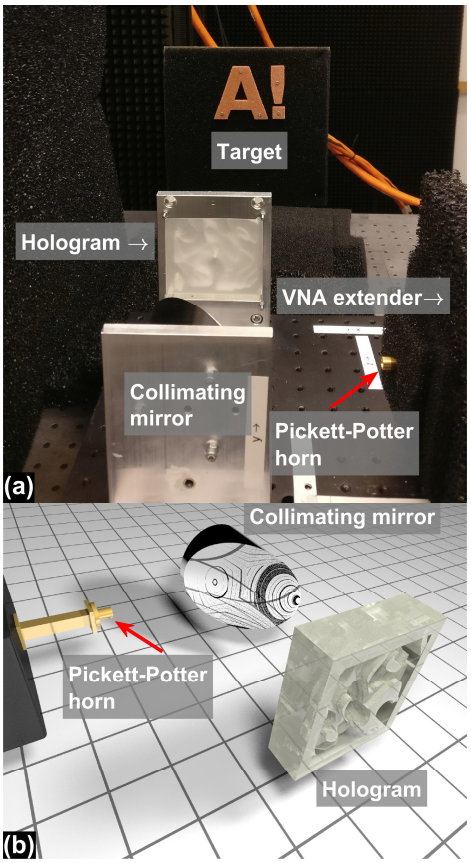

Fig. 4. a) A photograph of the imaging setup. The hologram is in the center of the photograph facing the Aalto-logo target, Pickett-Potter horn antenna and VNA extender are on the right. The collimating mirror is in the foreground facing the hologram. b) A rendered image from the opposite side to the photograph.

III. Measurement Results

A. Training Data for Test Targets

The imaging experiments were started by collecting training data for the neural network, and the test objects were scanned in the region of interest in 300-mm × 300-mm grid at 50-75 GHz and in 200-mm × 200-mm grid at 220-330 GHz. Scattering parameter S_{11} was measured when the Aalto logos were positioned at 10,201 locations in a 101 × 101 raster scan. The raster scan is not required in the imaging process, but it is used to deliver the training data (10,201 + 10,201 measurements) for the neural network. Once trained with vast amount of measurement data, only single measurement without mechanical scanning is used in imaging.

The reflection from the region of interest was extracted from the measured S_{11} by subtracting average S_{11} and time gating. The reflections from the quasioptics were outside the target

location time gate and only the contributions from 6 to 8 ns in time-domain were passed through the gate. This time window approximately corresponds from 0.9 to 1.2 m in range, which captures the reflection starting from distance, which equals the path length through quasioptics (~300 mm) and from the hologram aperture to the region of interest (600 mm). Subtracting the average S_{11} removes the background reflection that appears in the gating window for all test-object locations. Fig. 5 shows the processed reflection from the region of interest when the logos were 1) in the corner, 2) in the middle, and 3) incrementally shifted from the middle position by 2 mm (WR-3.4) and 3 mm (WR-15). The wide-band reflections from the two logos are presented in complex plane in 1001 frequency points for the target locations 1-3. The reflection is at a level of −33…−35 dB at WR-15 and −28…−20 dB at WR-3.4. In the figures, the amplitude of the reflection is normalized for convenience.

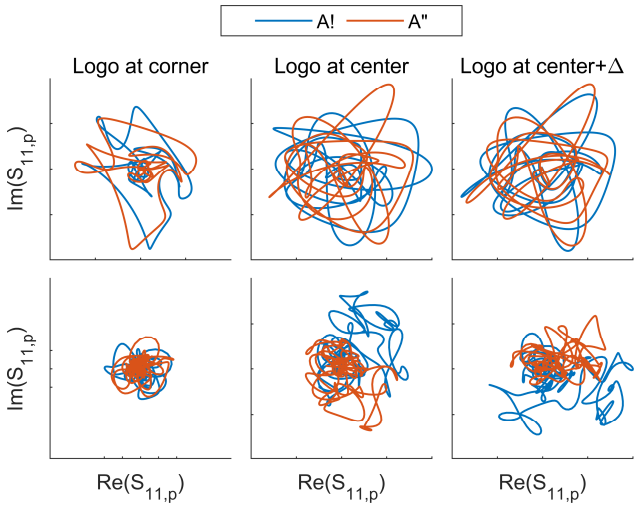

Fig. 5. Processed S_{11} measurements when Aalto logos A! and A" are at different locations in RoI. The top row show reflection at WR-15 and bottom row at WR-3.4.

Compared to the WR-15 measurements, the electrical sizes of the Aalto University logos are larger at WR-3.4 and the reflection loci in the polar plots include more detail. Aalto University logo in the corner of the scan area results in clearly distinct reflection when compared to middle position. For the higher-frequency band, the average amplitude drops more towards the corner position than in case of the lower band. This suggests more scattering at lower frequency and more directivity at higher frequency. The differences between loci at the incrementally separate positions are small for low-frequency and more significant at higher frequency – a qualitative indication of lower resolution at longer wavelength. At WR-15 the incremental translation is 0.6 free-space wavelengths at center frequency 62.5 GHz whereas at WR-3.4 it is 1.8 wavelengths at 275 GHz. It is evident that the characteristics of the test object are carried to the measured reflectivity: The reflection from logo A! has different spectral details than that from logo A". However, the function to map the reflection spectrum to an image is generally unknown.

B. *Augmented Data*

Measurements covering a large parameter space is challenging to achieve with physical test object as shown in Fig. 4(a): The object can only be replaced manually, which is not practical for many different test objects needed. In order to effectively train the neural network for a larger parameter space, data augmentation with a corner-cube reflector was carried out. The reflection from corner cube is measured in a 200-mm × 200-mm grid with a 2-mm scan step at WR-3.4. The corner cube diameter is 12.7 mm, and the ground truth image of it is a disk of the same size. The MNIST dataset of hand-written digits was used as the source of the data augmentation. An algorithm carries an overlap of the binarized 128 × 128 images of the digits. Each pixel values of the corner cube images were added in the augmented image if the corner cube overlaps with the MNIST binarized image of a digit. When the cube locations overlap with the digit image, the corresponding reflections from the cube are cumulatively added to form the augmented input to the neural network. Fig. 6 shows the augmented images and their augmented reflections.

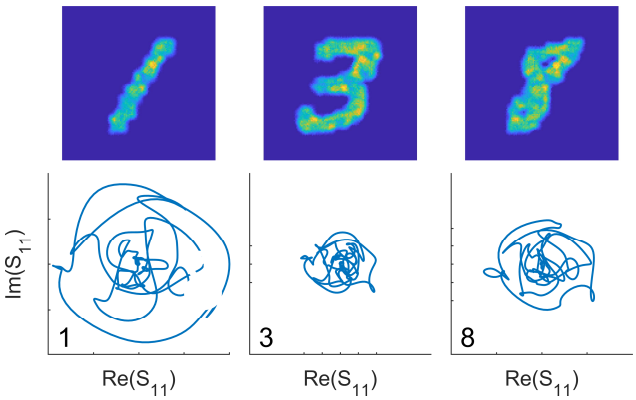

Fig. 6. Top row: Augmented images of the MNIST digits. The augmented images are summations of corner-cube images on the digit. Bottom row: The augmented reflections summed from all corner-cube locations on the digits.

IV. Deconvolutional Neural Network

The measured reflection was mapped into an image through a deconvolutional (DC) neural network. In deconvolution, reflection at each frequency at the input of the neural network is gradually spread across the image (the region of interest) at the output of the network. Reflection at each frequency is weighted with a collection of deconvolution kernels starting from each frequency component and increasing the layer dimensions through the network to arrive at the final image size, 256 × 256 pixels. Once the suitable weights were trained in the neural network, the reflectivity map of the region of interest was resolved from the spectra shown in Fig. 5. Our hypothesis is that the neural network can be trained to learn such generalizations that eventually a large parameter space (3D region of interest with multiple object types) can be covered. The ability for neural network to map from 1D to 3D computer graphics has been demonstrated with convolutional neural networks in [39]. Deconvolutional neural networks are better suited for image reconstruction, as is in our case too [23]. The

topology of the used deconvolutional neural network is shown in Fig. 7.

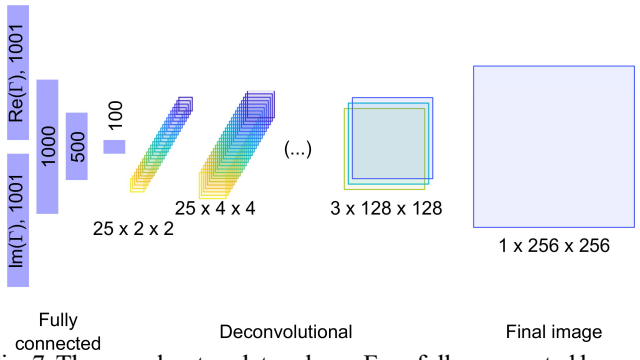

Fig. 7. The neural-network topology: Four fully-connected layers decrease the input to a latent representation with 100 neurons. Seven (four shown) deconvolutional layers expand the latent representation into an image.

The neural networks were trained separately for 50-75 GHz and 220-330 GHz reflectivity datasets. The input data to the neural network was the complex reflection with its real and imaginary parts concatenated to a 2×1001 long vectors. The network input data was min-max normalized to the interval [-1, 1]. In order to verify the learning of the latent representations, a fully-connected neural network reshapes the spectral content to a latent layer with only 100 neurons before the deconvolutional layers. From the latent layer onwards, the activations are reshaped to a deconvolutional neural network with increasing dimensions in successive layers. The deconvolutional part of the neural network with input reflection from Aalto logo A" is shown in Fig. 8. The network is symmetric so that all the parameters are the same for the both dimensions in the 2D layers. The output size of one dimension in a deconvolutional layer is

$$o' = s(i' - 1) + k, \tag{7}$$

where $s = 2$ is stride, i' is input layer size, and $k = 3$ is kernel size [40]. For example, the output size of the first deconvolution is given by $2(2 - 1) + 3 = 5$. The padding in the layers was selected to be "same", to have the 1st deconvolution output size of $4 = 2^2$ and to arrive to $2^8 = 256$ pixels per dimension at the network output. The neural network was developed in the Keras machine-learning API [41]. The neural network was trained by minimizing the mean squared error of the ground truth and the predicted image. The training utilized 2 × 10000 positions of the two Aalto logos across the region of interest. The predictions were evaluated with the logos at the same middle position in x-axis but varying in y-axis in 201 positions not shown to the neural network before.

The details in the RoI are build up from the small sub-images with increasing detail. Checkerboard-like details are carried out through the network. Some artifacts also appear in the final layer output. The checkerboard-like patterns are due to parameters in the deconvolutions, s and k, and are compensated in the final image to some extent through the learning [42]. To limit the appearance of the artifacts and to improve the convergence in network training, the image size could be tailored to the deconvolution parameters so that no overlap exists. However, in this case, the overlap was compromised over the demand to arrive in image size corresponding to the labels (256 × 256).

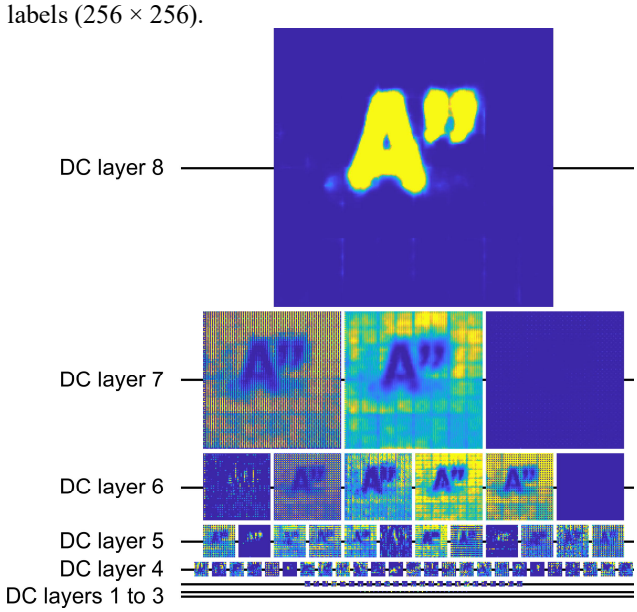

Fig. 8. Feature map weights of the deconvolutional layers in the neural network. Raster-like pattern is characteristic for deconvolution, but it is largely compensated in the final image.

V. Imaging Results

A. Test-Target Images

Fig. 9(a,c) and Fig. 10(a,c) show 12 predicted images from datasets at WR-15 (454 mm × 400 mm image size) and WR-3.4 (354 mm × 300 mm image size), respectively. Fig. 9(b,d) and Fig. 10(b,d) show the absolute error in the images. The predicted images were ranked by the mean squared sum of pixel-value error (MSE). The predictions with smallest error are high resolution images demonstrating very good agreement with the ground truth while the predictions with largest error may bear some similarities to the labels but are often incorrectly located. The average MSE is 0.099 at 50-75 GHz and 0.049 at 220-330 GHz. At lower frequency the MSE distribution is more uniform than at higher frequency, where the vast majority of the predictions is close to small error. The edge-spread function (ESF) in the predicted images is estimated across the edge at the top of letter "A". The ESF is determined for the 50-% percentile of the predictions with the least MSE. At WR-15, the FWHM of ESF is 15.0 mm while it is 10.5 mm at WR-3.4. MSE and ESF are shown in Fig. 11.

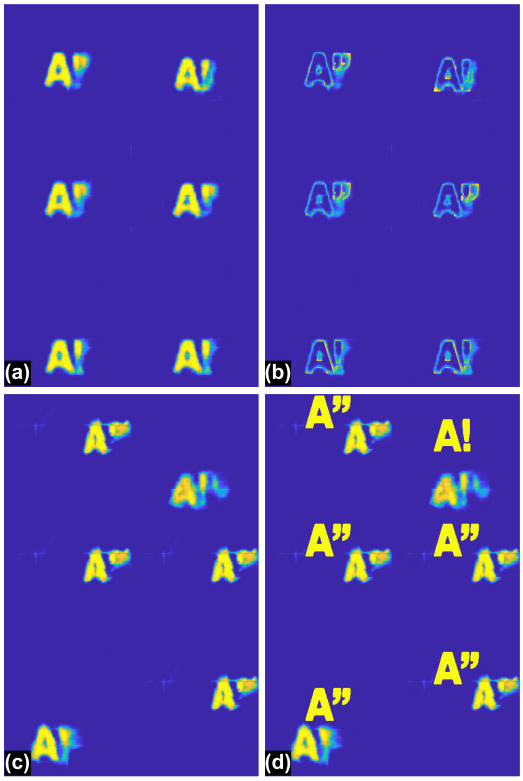

Fig. 9. a) Images and b) absolute value of image error for predictions with the least MSE at WR-15. The c) images and d) error with the greatest MSE.

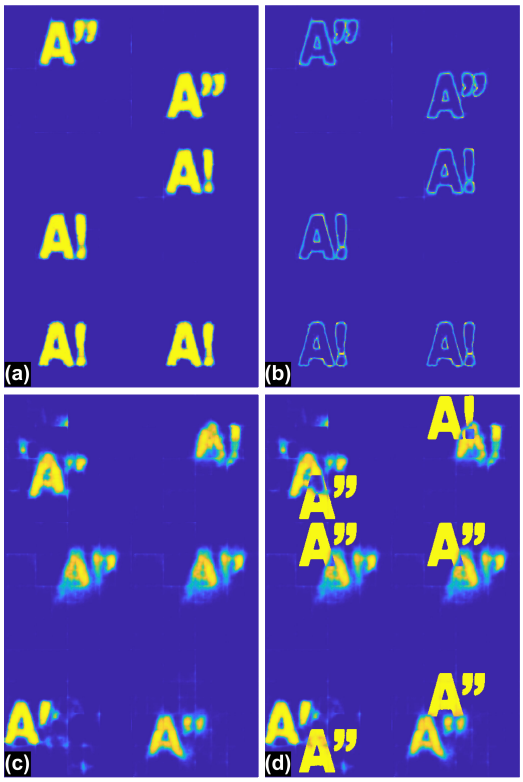

Fig. 10. a) Images and b) absolute value of image error for predictions with the least MSE at WR-3.4. The c) images and d) error with the greatest MSE.

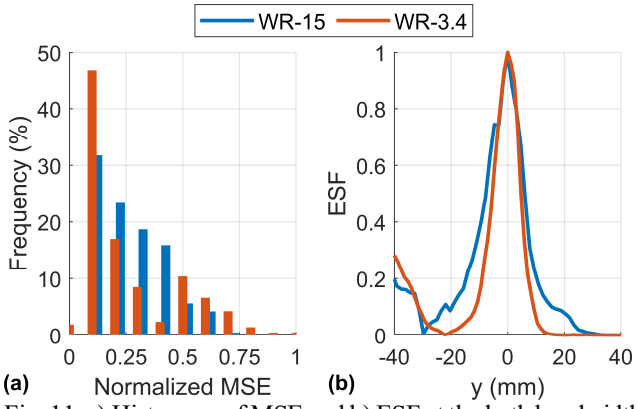

Fig. 11. a) Histogram of MSE and b) ESF at the both bandwidths.

B. Loss and Accuracy

In neural networks, successful predictions may be the result of overfitting. Then, the underlying abstractions are not learnt rather the correct prediction is memorized through heavily forcing the limited number of input data to the correct predictions. Overfitting may occur when the neural network is overly complex and certain input activations are emphasized over others for, e.g., a location of the test object. To counter overfitting, we use the fully-connected layer of latent representation (100 neurons), dropout layer with 50-% dropout probability, and limited complexity of the neural network. The neural networks have ~ 2.6 million degrees of freedom, and they were trained with mean squared error as the loss function. Training was evaluated at the end of each epoch and there were 500 epochs in total for both datasets. The loss was evaluated for both the training and validation datasets. The data in validation is not given in training, and hence validation loss converging simultaneously with the training loss is a sign of learning. Fig. 12 shows the mean absolute error (MAE) and loss for the datasets during training and validation.

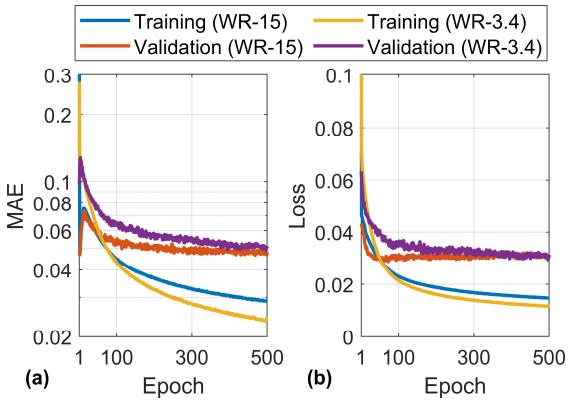

Fig. 12. a) Mean absolute value of error and b) loss for the WR-15 and WR-3.4 datasets.

The WR-15 and WR-3.4 datasets have different baselines due to the image size differences. Both data sets converge fast at the beginning of the iteration. The WR-3.4 dataset shows training and validation that both converge up to epoch 500. In case of WR-15 dataset, the convergence differentiates between the training and validating, suggesting some overfitting in training that does not convey to validation (learning occurs less). Overfitting occurs when the training continues to converge

while validation remains unchanged. At WR-3.4, the slope in validation loss is negative (learning occurs) whereas for WR-15 the slope is close to zero. To conclude: At WR-3.4, while some overfitting may exist, the neural network was able to learn generalizations in the described setting and form high-resolution imagery.

C. *Augmented Images*

Fig. 13 shows the augmented images for WR-3.4. The successful prediction of augmented images is an indication that overfitting does not occur to a great extent. Further, the PSF with augmented data should be similar to the data from extended object to justify its use. For data-processing purposes the image size is 128 × 128 pixels. The neural network was trained with 60000 augmented reflections from the MNIST digits (Fig. 6). As opposed to the truly experimental imaging tests with the Aalto University logos, the augmented data represents a much larger set of different labels and input data. Therefore, it was assumed that augmented data tests the robustness of such a neural network to learn generalizations of complex reflectivity patterns. In contrast from the classical MNIST digit recognition task, we used the hand-written digit examples as the label and augment the reflection from the digit-shaped object. The predictions in Fig. 13 were not shown to the neural network during the training, yet they agree with the labels with an MSE less than 0.02 at WR-3.4. In general, the predictions display smoother edges than the corresponding labels. Some of the labels have not simply connected regions, e.g., a separated dot. These details seem to merge to the digit or background in the image predictions.

As the labels in the augmented dataset are each different, the global point-spread function (PSF) is determined with a fitting method similar to the one presented in [43]. The PSF is approximated with a Gaussian function

$$g(a,\sigma) = a\exp(-(r/\sigma)^2), \tag{8}$$

where σ is the width of the Gaussian function in the cross-range coordinates and a is an amplitude factor. The PSF is found by minimizing

$$L(a,\sigma) = \sum_{x,y}\left(\hat{y} - \left(y * g(a,\sigma)\right)\right)^2, \tag{9}$$

where $\hat{y}$ is the prediction, y is the label, and $*$ is the convolution operator. The minimizing of L is equal to finding a Gaussian blur kernel that reduces the image to the same resolution as the prediction from the neural network. The resolution in the images can be approximated by the full width at half maximum (FWHM) value of the PSF.

The Gaussian function (8) relates to the resolution as $\mathrm{FWHM} = 2\sqrt{\ln 2}\,\sigma$. Fig. 14 shows the estimates of FWHM for the experimental and augmented datasets. The shown histograms verify that in general predictions at WR-15 have greater variance than at WR-3.4. In case for WR-3.4, the FWHM distribution is well centered at 16.5 mm, close to the theoretical value from 16 to 19 mm. For the case of the augmented data at WR-3.4, the FWHM is surprisingly narrow. This may be due to the averaging effects in data augmentation, which may bring the noise down.

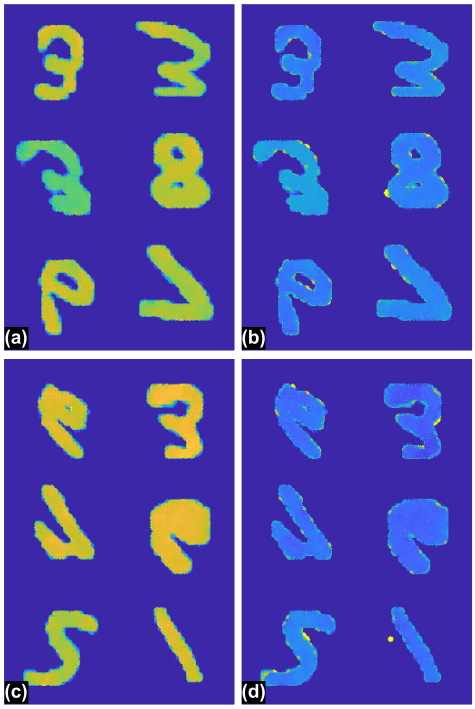

Fig. 13. a) Images and b) absolute value of image error for predictions with the least MSE for augmented dataset at WR-3.4. The c) images and d) error with the greatest MSE.

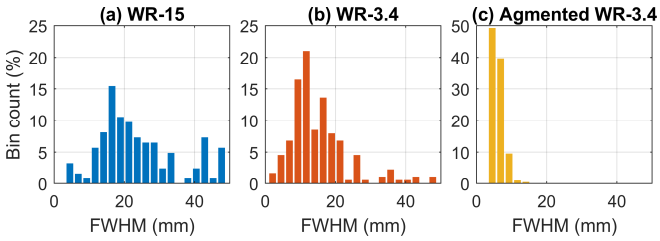

Fig. 14. Histograms of the FWHM with a) WR-15, b) WR-3.4, and c) augmented WR-3.4 datasets.

D. *Latent Representations*

The latent representations were studied for the datasets with Aalto logos and augmented data sets with MNIST digits, both at the WR-3.4 band. Fig. 15 shows the 3rd hidden layer outputs for 212 predictions with Aalto logos and for 200 predictions with augmented MNIST digits. The output is from a layer with 100 fully connected neurons, before the deconvolutional layers in the network. The layer output can be considered a latent representation of the test object. The horizontal axis represents each of the 100 neurons in the 3rd hidden layer (latent representation). The vertical axis represents different targets and locations, either Fig. 15(a) A! or A" increasing in *y*-axis location or Fig. 15(b) MNIST digit numerical value increasing with multiple samples of the same digit. The number of samples for each digit vary due to random nature of the dataset. The latent representations in Fig. 15 are sorted with Aalto logo *y*-axis location. The MNIST logos were sorted with increasing digit numerical value. The test objects with the closest resemblance displayed increased similarities in the latent representations.

The latent representations of the Aalto University logos are sparser compared to the MNIST representation: 32 to 35 % of

the neuron outputs are non-zero for Aalto logos, whereas 64 to 67 % are non-zero for MNIST augmented data. The fraction of non-zero pixel values in the final image are ~8 % for Aalto logo and ~19 % for MNIST images. The results may indicate that relatively few inputs are needed to map to two different test objects located in the region of interest whereas the MNIST consists greater number of mappings from reflection to image, hence requiring also more non-zero values in the latent representation. For both cases, limiting the number of values in latent representation to 100 appears sufficient with respect to image quality.

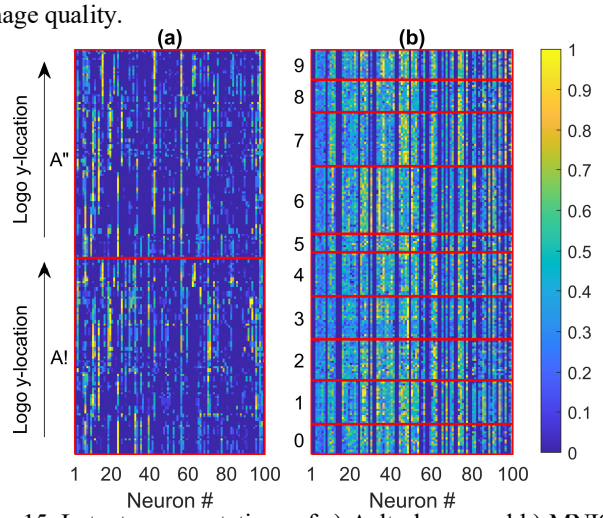

Fig. 15. Latent representations of a) Aalto logos and b) MNIST digits in the hidden layer in the neural network.

VI. Conclusions

In this work we have, for the first time, demonstrated millimeter- and submillimeter-wave imaging based on wide-band reflection measurements through a highly dispersive dielectric hologram. The hologram was designed to produce a quasirandom goal field in the region of interest at the WR-15 and WR-3.4 bands. The hologram produced multi-wavelength phase shifts across the two bandwidths so that the field at the region of interest is highly dispersive. Two experimental campaigns at 50-75 GHz and 220-330 GHz were carried out with highly scattering reflective targets. Two separate deep neural networks based on deconvolutional layers were trained for the bandwidths with the photographs from the region of the interest at a distance of 600 mm. The resulting image quality is in line with the diffraction limit as estimated via illumination frequency and hologram aperture. The experimental resolution estimated from the global point-spread function was ~17 mm at best at WR-3.4. The edge-spread function shows locally even slightly better resolution, which may indicate slight overfit of the neural network. Overfitting often occurring in neural networks was mitigated by use of dropout layer and limited number of neurons in the latent representation. The close-to-theoretical resolution and analysis of loss in the train and validation datasets during training indicate that the wave propagation and scattering are learnt.

This initial work was carried out with a limited number of real test objects, and data with significantly more test objects are needed to further learn generalizations. The augmented dataset with 60000 examples of MNIST digits presented a vast collection of different types of test objects. The image predictions from the augmented reflection from MNIST data suggest that the presented imaging methodology may be able to image more complex test objects too. A large percentage of the latent representation values are close to zero, which suggests that 100 neurons suffice to form an image in the cases presented here.

The presented method enables efficient computational millimeter- and submillimeter-wave imaging without the need for optomechanically scanned systems or complex phase-shifter or transceiver arrays. Currently, the field of view and imaging distance are limited to the hologram aperture and illumination edge taper, but the research so far foresees the possibility to scale up to, e.g., personnel screening application.

Acknowledgements

This work has been carried out in MilliLab, an external laboratory of the European Space Agency. Mr. Markus Grönholm from Alshain Oy is thanked for the hologram manufacture.

References

[1] S. S. Ahmed, “Microwave Imaging in Security — Two Decades of Innovation,” *IEEE Journal of Microwaves*, vol. 1, no. 1, pp. 191-201, 2021.

[2] Heinz, T. May, D. Born, G. Zieger, S. Anders, V. Zakosarenko, H. -. Meyer and C. Schäffel, “Passive 350 GHz Video Imaging Systems for Security Applications,” *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 36, no. 10, pp. 879-895, 2015.

[3] J. Luomahaara, H. Sipola, L. Grönberg, A. Mäyrä, M. Aikio, A. Timofeev, K. Tappura, A. Rautiainen, A. Tamminen, V. Vesterinen, M. Leivo, F. Gao, H. Vasama, A. Luukanen and J. Hassel, “A Passive, Fully Staring THz Video Camera Based on Kinetic Inductance Bolometer Arrays,” *IEEE Transactions on Terahertz Science and Technology*, vol. 11, no. 1, pp. 101-108, 2021.

[4] D. A. Robertson, D. G. Macfarlane, R. I. Hunter, E. Gandini and N. Llombart, “Real beam, volumetric radar imaging at 340 GHz for security applications,” *Proceedings of the IEEE International Symposium on Antennas and Propagation and USNC/URSI National Radio Science Meeting*, 2018, pp. 615-616.

[5] S. S. Ahmed, A. Genghammer, A. Schiessl and L. Schmidt, “Fully Electronic E-Band Personnel Imager of 2 m^2 Aperture Based on a Multistatic Architecture,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 1, pp. 651-657, 2013.

[6] S. S. Ahmed, A. Schiessl and L. Schmidt, “A Novel Fully Electronic Active Real-Time Imager Based on a Planar Multistatic Sparse Array,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 59, no. 12, pp. 3567-3576, 2011.

[7] S. Augustin, S. Frohmann, P. Jung and H. Hübers, “Mask Responses for Single-Pixel Terahertz Imaging,” *Scientific Reports*, vol. 8, no. 1, pp. 4886, 2018.

[8] R. I. Stantchev, D. B. Phillips, P. Hobson, S. M. Hornett, M. J. Padgett and E. Hendry, “Compressed sensing with near-field THz radiation,” *Optica*, vol. 4, no. 8, pp. 989-992, 2017.

[9] C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith and W. J. Padilla, “Terahertz compressive imaging with metamaterial spatial light modulators,” *Nature Photonics*, vol. 8, no. 8, pp. 605-609, 2014.

[10] R. I. Stantchev, X. Yu, T. Blu and E. Pickwell-MacPherson, "Real-time terahertz imaging with a single-pixel detector," *Nature Communications*, vol. 11, article no. 2535, 2020.

[11] D. M. Sheen, D. L. McMakin and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 9, pp. 1581-1592, 2001.

[12] L. Qiao, Y. Wang, Z. Zhao and Z. Chen, "Range Resolution Enhancement for Three-Dimensional Millimeter-Wave Holographic Imaging," *IEEE Antennas and Wireless Propagation Letters*, vol. 15, pp. 1422-1425, 2016.

[13] T. Zvolensky, V. R. Gowda, J. Gollub, D. L. Marks and D. R. Smith, "W-Band Sparse Imaging System Using Frequency Diverse Cavity-Fed Metasurface Antennas," *IEEE Access*, vol. 6, pp. 73659-73668, 2018.

[14] Y. Yu, L. Qiao, Y. Wang and Z. Zhao, "Active millimeter wave three-dimensional scan real-time imaging mechanism with a line antenna array," arXiv Preprint arXiv:2102.04878, 2021.

[15] S. Venkatesh, X. Lu, H. Saeidi and K. Sengupta, "A high-speed programmable and scalable terahertz holographic metasurface based on tiled CMOS chips," *Nature Electronics*, vol. 3, no. 12, pp. 785-793, 2020.

[16] T. F. Gallacher, D. A. Robertson and G. M. Smith, "The Photo-Injected Fresnel Zone Plate Antenna: Optoelectronic Beam Steering at mm-Wave Frequencies," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1688-1696, 2013.

[17] T. F. Gallacher, R. Sondena, D. A. Robertson and G. M. Smith, "Optical Modulation of Millimeter-Wave Beams Using a Semiconductor Substrate," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 7, pp. 2301-2309, 2012.

[18] M. K. Leino, J. Bergman, J. Ala-Laurinaho and V. Viikari, "Millimeter-Wave Frequency-Diverse Imaging with Phased Array Intended for Communications," *Progress in Electromagnetics Research M*, vol. 101, pp. 69-78, 2021.

[19] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications*, vol. 7, article no. 17141, 2018.

[20] A. Sinha, J. Lee, S. Li and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117-1125, 2017.

[21] Mousavi and R. G. Baraniuk, "Learning to invert: Signal recovery via deep convolutional networks," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 2272-2276.

[22] M. D. Zeiler, G. W. Taylor and R. Fergus, "Adaptive deconvolutional networks for mid and high level feature learning," *Proceedings of the International Conference on Computer Vision*, 2011, pp. 2018-2025.

[23] H. Noh, S. Hong and B. Han, "Learning deconvolution network for semantic segmentation," *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1520-1528.

[24] M. ELsaadouny, J. Barowski and I. Rolfes, "Non-destructive testing of 3D-printed samples based on machine learning," *Proceedings of the IEEE MTT-S International Microwave Workshop Series on Advanced Materials and Processes for RF and THz Applications (IMWS-AMP)*, 2019, pp. 22-24.

[25] H. Park and J. Son, "Machine Learning Techniques for THz Imaging and Time-Domain Spectroscopy," *Sensors (Basel, Switzerland)*, vol. 21, no. 4, article no. 1186, 2021.

[26] D. Liao, K. F. Chan, C. H. Chan, Q. Zhang and H. Wang, "All-optical diffractive neural networked terahertz hologram," *Optics Letters*, vol. 45, no. 10, pp. 2906-2909, 2020.

[27] Y. Luo, D. Mengu, N. T. Yardimci, Y. Rivenson, M. Veli, M. Jarrahi and A. Ozcan, "Design of task-specific optical systems using broadband diffractive neural networks," *Light: Science & Applications*, vol. 8, no. 1, article no. 112, 2019.

[28] J. Li, D. Mengu, N. T. Yardimci, Y. Luo, X. Li, M. Veli, Y. Rivenson, M. Jarrahi and A. Ozcan, "Spectrally encoded single-pixel machine vision using diffractive networks," *Science Advances*, vol. 7, no. 13, article no. eabd7690, 2021.

[29] J. Li, Y. Li, Y. Cen, C. Zhang, T. Luo and D. Yang, "Applications of Neural Networks for Spectrum Prediction and Inverse Design in the Terahertz Band," *IEEE Photonics Journal*, vol. 12, no. 5, pp. 1-9, 2020.

[30] Y. Li, W. Hu, X. Zhang, Z. Xu, J. Ni and L. P. Ligthart, "Adaptive terahertz image super-resolution with adjustable convolutional neural network," *Optics Express*, vol. 28, no. 15, pp. 22200-22217, 2020.

[31] S. Agarwal, A. S. Bisht, D. Singh and N. P. Pathak, "A novel neural network based image reconstruction model with scale and rotation invariance for target identification and classification for Active millimetre wave imaging," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 35, no. 12, pp. 1045-1067, 2014.

[32] A. Tamminen, S. Pälli, J. Ala-Laurinaho, A. Aspelin, A. Oinaanoja and Z. Taylor, "Holograms with neural-network backend for submillimeter-wave beamforming applications," *Proceedings of the SPIE: Passive and Active Millimeter-Wave Imaging XXIII*, 2020, pp. 114110C.

[33] M. D. Zeiler, D. Krishnan, G. W. Taylor and R. Fergus, "Deconvolutional networks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2528-2535.

[34] K. B. Cooper, R. J. Dengler, N. Llombart, T. Bryllert, G. Chattopadhyay, E. Schlecht, J. Gill, C. Lee, A. Skalare, I. Mehdi and P. H. Siegel, "Penetrating 3-D Imaging at 4- and 25-m Range Using a Submillimeter-Wave Radar," *IEEE Transactions on Microwave Theory and Techniques*, vol. 56, no. 12, pp. 2771-2778, 2008.

[35] S. -V. Pälli, A. Tamminen, J. Ala-Laurinaho and Z. D. Taylor, "Design and Characterization of Phase Holograms for Standoff Localization at Millimeter and Submillimeter Waves," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 1, pp. 907-918, 2022.

[36] Goldsmith, Paul F. *Quasioptical systems*. Chapman & Hall, 1998.

[37] S. -V. Pälli, A. Tamminen, J. Ala-Laurinaho and Z. Taylor, "Dielectric phase hologram for frequency-diverse millimeter and submillimeter-wave imaging applications," *Proceedings of the Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)*, 2021, paper cc_5_5.

[38] J. W. Lamb, "Miscellaneous data on materials for millimetre and submillimetre optics," *International Journal of Infrared and Millimeter Waves*, vol. 17, no. 12, pp. 1997-2034, 1996.

[39] J. Wu, C. Zhang, T. Xue, W. T. Freeman and J. B. Tenenbaum, "Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling," arXiv Preprint arXiv:1610.07584, 2016.

[40] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," arXiv Preprint arXiv:1603.07285, 2016.

[41] Keras: the Python deep learning API. Available: https://keras.io/.

[42] A. Odena, V. Dumoulin and C. Olah, "Deconvolution and checkerboard artifacts," *Distill*, vol. 1, no. 10, article no. e3, 2016.

[43] N. Joshi, R. Szeliski and D. J. Kriegman, "PSF estimation using sharp edge prediction," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1-8.

Aleksi Tamminen was born in Ruotsinpyhtää, Finland, in 1982. He received the B.Sc. (Tech.) and M.Sc. (Tech.) degrees in electrical engineering from the Helsinki University of Technology, Espoo, Finland, in 2005 and 2007, respectively. He received the Lic.Sc. (Tech.) and D.Sc. (Tech.) degrees from Aalto University (former Helsinki University of Technology), Espoo, in 2011 and 2013, respectively. From 2005 to 2013, he was with the Department of Radio Science and Engineering, Aalto University. His research work in Aalto University was related to antenna measurements and imaging at millimeter and submillimeter waves. From 2013 to 2018 he was Research Scientist at Asqella Ltd., Helsinki, Finland. He was the principal in the research and development of commercial submillimeter-wave imaging technology as well as in participating in academic research projects in the related field. From 2018, he is with Aalto University as Research Fellow. His current research interests are submillimeter- and millimeter-wave projects including antenna measurements, sensing biological tissues, and imaging.

Samu-Ville Pälli was born in Savonlinna, Finland, in 1994. He received the B.Sc. (Tech.) (Hons.) and M.Sc. (Tech.) (Hons.) degrees in electrical engineering, in 2018 and 2020, respectively, from Aalto University, Espoo, Finland, where he is currently pursuing the D.Sc. (Tech.) degree. He has been with the Department of Electronics and Nanoengineering, School of Electrical Engineering, Aalto University, since 2017. His current research interests include computational imaging methods at submillimeter waves and THz imaging and sensing.

Juha Ala-Laurinaho received the Diploma Engineer (M.Sc.) degree in mathematics and D.Sc. (Tech.) degree in electrical engineering from TKK Helsinki University of Technology, Finland, in 1995 and 2001, respectively. He has been with the TKK, currently Aalto University, serving in the Radio Laboratory in 1995–2007, in the Department of Radio Science and Engineering in 2008-2016, and currently in the Department of Electronics and Nanoengineering. Currently, he works as a Staff Scientist. Dr. Ala-Laurinaho has been a Researcher and Project Manager in many millimeter wave technology related projects. His current research interests are the antennas and antenna measurement techniques for millimeter and submillimeter waves, and the millimeter wave imaging.

Zachary D. Taylor (S'06–M'09) received the B.S. degree in electrical engineering from the University of California (UCLA), Los Angeles, CA, USA, in 2004, and the M.S. and Ph.D. degrees in electrical engineering from the University of California (UCSB), Santa Barbara, CA, USA, in 2006 and 2009, respectively. From 2013 – 2018 he was an adjunct assistant professor with appointments in the Department of Bioengineering, Department of Electrical Engineering, and Department of Surgery at the University of California (UCLA), Los Angeles, CA, USA. From 2018, he is an assistant professor in the Department of Electronics and Nanoengineering at Aalto University, Espoo, Finland. His current research interests include submillimeter-wave and THz imaging, sensing, and calibration.