White Rose Research Online
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/183016/

Version: Accepted Version

Article:

Eckels, J.D., Jacobson, E.M., Cummings, I.T. et al. (2022) Predicting local material thickness from steady-state ultrasonic wavefield measurements using a convolutional neural network. Ultrasonics, 123. 106661. ISSN: 0041-624X

https://doi.org/10.1016/j.ultras.2021.106661

© 2021 Published by Elsevier B.V. This is an author produced version of a paper subsequently published in Ultrasonics. Uploaded in accordance with the publisher's self-archiving policy. Article available under the terms of the CC-BY-NC-ND licence (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Predicting Local Material Thickness from Steady-State Ultrasonic Wavefield Measurements Using a Convolutional Neural Network

Joshua D. Eckels[a], Erica M. Jacobson[a], Ian T. Cummings[a,b], Isabel F. Fernandez[a], Kelly Ho[a], Nikolaos Dervilis[c], Eric B. Flynn[b], Adam J. Wachtor[a]

[a]*Engineering Institute, Los Alamos National Laboratory, Los Alamos, NM 87545*
[b]*Space and Remote Sensing, Los Alamos National Laboratory, Los Alamos, NM 87545*
[c]*Mechanical Engineering, The University of Sheffield, Sheffield, UK S1 3JD*

Abstract

Acoustic steady-state excitation spatial spectroscopy (ASSESS) is a full-field, ultrasonic non-destructive evaluation (NDE) technique used to locate and characterize defects in plate-like structures. ASSESS generates a steady-state, single-tone ultrasonic excitation in a structure and a scanning laser Doppler Vibrometer (LDV) measures the resulting full-field surface velocity response. Traditional processing techniques for ASSESS data rely on wavenumber domain analysis. This paper presents the alternative use of a convolutional neural network (CNN), trained using simulated ASSESS data, to predict the local plate thickness at every pixel in the wavefield measurement directly. The defect detection accuracy of CNN-based thickness predictions are shown to improve for defects of greater size, and for defects with higher thickness reductions. The CNN demonstrates the ability to predict thickness accurately in regions where Lamb wave dispersion relations are complex or unknown, such as near the boundaries of a test specimen, so long as the CNN is trained on data that accounts for these regions. The CNN also shows generalizability to ASSESS experimental data, despite an entirely simulated training dataset.

Keywords: acoustic steady-state excitation spatial spectroscopy, convolutional neural network, image segmentation, ultrasonic measurement

---

1. Introduction

Ultrasonic full wavefield imaging provides a means of detecting defects over large inspection areas [1]. Laser Doppler vibrometers (LDVs) are often used to make these full-field ultrasonic measurements, and a great deal of work has gone into developing processing techniques for the datasets they generate. For example, Ruzzene presented a method for identifying damage in structures in a transient ultrasonic scanning LDV system by isolating the excitation of the structure from the response caused by reflections from defects in the wavenumber domain [2]. Flynn and Jarmer successfully demonstrated the detection of hidden corrosion on the backside of an aluminum plate and delaminations in several composite plates via wavenumber imaging [3]. Rogge and Leckey demonstrated the effectiveness of wavenumber analysis at revealing disbonds in aluminum and composite plates when interrogated with transient ultrasonic signals and imaged via scanning LDV or

---

Email addresses: `eckelsjd@lanl.gov` (Joshua D. Eckels), `ejacobson@lanl.gov` (Erica M. Jacobson), `itc@lanl.gov` (Ian T. Cummings), `n.dervilis@sheffield.ac.uk` (Nikolaos Dervilis), `eflynn@lanl.gov` (Eric B. Flynn), `ajw@lanl.gov` (Adam J. Wachtor)

ultrasonic immersion C-scan [4]. Huthwaite and Simonetti described a transient laser ultrasonic iterative tomography based thickness estimation approach that improved thickness estimation accuracy over more traditional tomographic methods [5]. Aryan *et al.* described a system that used three LDVs to capture the full 3D surface displacement in response to transient excitation, and utilized the root mean square (RMS) amplitude of the excited waveform at each point as a feature to detect several kinds of damage [6].

Acoustic steady-state excitation spatial spectroscopy (ASSESS) is an inspection technique in which the test article is excited with a single ultrasonic tone and a scanning LDV records the full-field surface velocity response of the test article. ASSESS was first described by Flynn and Jarmer [3], where their key insight was exciting the structure using a single, steady-state ultrasonic tone. This allowed them to continuously scan the structure without having to transiently pulse the structure for each measurement and wait for the excitation to die out before moving to the next measurement. Furthermore, the steady-state nature of the excitation and the concentration of the amplifier power in a narrow frequency range significantly improved the signal-to-noise ratio (SNR) of the ASSESS measurements over transient measurement techniques.

Several other ASSESS-related works have also exploited wavenumber analysis for defect detection [7, 8, 9]. O'Dowd *et al.* explored the limits of wavenumber estimation using simulated ASSESS data [10]. Mesnil *et al.* described transient and steady-state wavenumber and thickness estimation systems for composite plates based on scanning LDVs [11] — the steady-state version was very similar to the one proposed in [3], except that they used phase gradients to estimate the local wavenumber and related the estimated wavenumber to thickness (composite ply count) via simulated Lamb wave dispersion curves. The thickness estimation procedure reported in [11] also included a refinement step to improve the quality of their results based on the Smooth-Sparse Decomposition.

As mentioned above, one of the valuable data products that can be extracted from ASSESS datasets is an estimated local thickness of the structure — enabling quantification of hidden corrosion and estimation of the depth of delamination, among other possibilities. Stull *et al.* presented theoretical upper bounds on thickness estimation performance in a number of ASSESS scenarios [12]. While the wavenumber and thickness estimation techniques described above appear to perform well in continuous regions of the plate, their results are difficult to interpret near boundaries and transition regions, where the Lamb wave dispersion relationships do not apply. This work presents a thickness estimation technique for ASSESS data based on a convolutional neural network (CNN) that shows accurate thickness predictions without relying on the Lamb wave dispersion relationships, potentially making it applicable to more complex geometric regions. This paper is an extension of Eckels *et al.* [13] to report new simulated and experimental results and performance parameter studies.

A number of papers have been written on processing ultrasonic NDE data with deep learning techniques. For example, Virkkunen *et al.* demonstrated crack detection in ultrasonic phased array data using deep CNNs [14]. Li *et al.* described a method for estimating one-dimensional material thickness profiles from measured displacements by first reconstructing them with a wavenumber spatial transform and then improving those estimates by processing them with a CNN [15]. Meng *et al.* successfully identified artificial delaminations and voids in a composite panel by feeding wavelet coefficients from immersion C-scans using a CNN with a linear support vector machine output layer that indicated the layer of the delamination/void or the absence of defect at each scan point [16]. Ijjeh *et al.* demonstrated a full-wavefield delamination detection scheme based on a fully convolutional neural network (FCN) that input RMS wavefield amplitudes from LDV response measurements from transient excitation and produced a binary mask indicating the presence or absence of delaminations [17]. Song and Yang demonstrated a two-stage super-resolution technique for detecting and quantifying the extent of subwavelength defects in aluminum plates [18]. The method created a virtual phased array with a scanning LDV and transient pulse excitation of the plate with a piezoelectric transducer (PZT) patch. An initial image was formed using traditional delay-and-sum beamforming and was then processed by a CNN to establish a probabilistic defect detection mask, which was overlaid on the original image to isolate the defects from the noise. Small segments containing defects were then reanalyzed with a second CNN to more accurately resolve the spatial extent of the defects.

The work presented in this paper differs in several important regards from the work mentioned above. The ASSESS scan technique is inherently narrowband and uses a continuous excitation, making the scan rate many times faster than typical transient ultrasonic LDV scans (i.e. seconds or minutes per scan instead of hours). The decreased measurement time is caused by the continuous excitation significantly improving the SNR of the scans through continuous energy transfer into the structure and avoiding the need for averaging several measurements at a given pixel. Our technique is designed to directly estimate plate thickness based on wavefield images at the excitation frequency. The only pre-processing that is performed on the raw LDV signal is what is necessary to convert it into the complex, out-of-plane, surface velocity response wavefield of the test article. The ideal, harmonic steady state response to an excitation at frequency ω_E at a spatial coordinate (x, y) and time t, as measured using the LDV, takes the form

$$v\left(x,y,t\right)=a\left(x,y\right)\cos\left(\omega_{E}t\right)+b\left(x,y\right)\sin\left(\omega_{E}t\right). \tag{1}$$

where a and b are the real and imaginary coefficients, respectively. During the ASSESS measurement process, this measured response is converted to a compact, complex representation r using the Fourier transform at the excitation frequency

$$r\left(x,y\right)=\int v\left(x,y,t\right)e^{-i\omega_{E}t}dt\approx a\left(x,y\right)+ib\left(x,y\right), \tag{2}$$

where i is the complex unit. The real and imaginary parts can be thought of as two snapshots of the steady-state response, which can be used to recreate the measured response at any point in the harmonic cycle, $\theta \in [0, 2\pi]$, through

$$v\left(x,y,\theta\right)=a\left(x,y\right)\cos\left(\theta\right)+b\left(x,y\right)\sin\left(\theta\right). \tag{3}$$

It is important to note that the CNN model was trained using simulated data because gathering enough experimental data to train the model would be prohibitively expensive. We show that the model's performance generalizes reasonably well to experimental data.

The remaining sections of the paper are outlined as follows. Section 2 describes details of the CNN model, parameters of the simulated data used to train and test the CNN model, and metrics used to evaluate the model performance. New simulated test cases beyond those reported in [13] and an extension of the CNN model to thickness predictions of experimentally collected data sets are presented in Section 3. The results of these cases demonstrate that the CNN has learned a generalized relationship between the raw wavefield from an ASSESS measurement and the local thickness at each pixel in the inspected plate.

2. Materials and Methods

2.1. CNN architecture

Image segmentation is a well-known task in computer vision where images are classified on a pixel level. It can be thought of as an extension of typical object detection techniques, where instead of predicting bounding boxes of objects in an image, it highlights filled regional boundaries of class instances in an image. For more details on different deep learning frameworks used for image segmentation and object detection, see the following surveys: [19], [20], and [21]. A popular CNN architecture used for image segmentation is the U-Net segmentation model, proposed by Ronneberger et al. [22]. The U-Net is a type of FCN that uses a series of down-sampling convolutional layers in the encoder, followed by a series of up-sampling deconvolutional layers in the decoder, with horizontal cross-connections to carry feature maps over from the encoder to improve pixel-wise prediction in the output segmentation map. All network parameters in an FCN can be learned through standard neural network training techniques. The CNN used in this study is a U-Net style FCN with an encoder based on the 34-layer ResNet model proposed by He et al. [23]. The architecture implementation details are provided in the Fast.AI library [24].

The input to the CNN is a 400 × 400 pixel image of the real part of the full-field, steady-state complex surface response of a plate-like structure to a single-tone ultrasonic excitation. The real part of the complex response was arbitrarily chosen over the imaginary given that the two wavefield images are nearly indistinguishable. The expected output of the CNN is a 400 × 400 segmented image, where each pixel is a prediction of the local thickness of the plate. The classification at each pixel will be a discrete integer value from $1 - 10$ mm, since this is a practical range for most industrial applications of thin aluminum plates. The CNN is also provided ground truth segmentation images to use as labeled data during supervised learning. After passing an input image through the network, it passes through a softmax function, where the CNN outputs a tensor of size 10 × 400 × 400. The first dimension of the tensor contains class probability values at each pixel in the 400 × 400 output image; the size of this dimension is equal to the number of thickness classes (i.e. 10). The final predicted segmentation image is obtained by taking the maximum probability along the first dimension of the raw output.

2.2. Simulation technique

A persistent problem in the field of deep neural networks is access to substantial training datasets. Because experimental ASSESS data is not readily available in sufficient quantities, a simulation procedure was used instead to gather a large enough dataset. The simulation procedure outlined here follows a previous study by O'Dowd et al., which used a three-dimensional finite element analysis (FEA) to simulate the ASSESS experimental procedure in ANSYS [10]. A thin, 400 mm × 400 mm aluminum alloy plate was modeled in CAD software, loaded into a harmonic response analysis in ANSYS, and meshed using tetrahedrons. The edges of the plate were set as free boundary conditions and to emulate PZT excitation, a small, 40 mm diameter extrusion was placed on the surface of the plate and excited by a 0.1 MPa pressure at a frequency of 80 kHz. Although a rectilinear mesh could have been used for the simple plate geometry, tetrahedrons provided a better geometric representation for simulations containing curved defect shapes. This simulation procedure is a simplification of the true contact condition that exists between the plate and the PZT, but was sufficient for the current study where capturing the wavefield in large areas of the plate was of primary concern. O'Dowd et al. used the same simulation procedure and reported convergence for a mesh size of 2 mm, which was used in all simulations performed for this study.

The complex, steady-state, out-of-plane deformation response of the plate emulates the experimental data gathered by an LDV during ASSESS experiments (see Eq. (2)). The real component of the deformation response of all external mesh nodes in the simulation was exported to MATLAB for further processing. The response of the top surface nodes was linearly interpolated to a uniform 400 × 400 grid and saved as a grayscale wavefield image, with the amplitude scaled from zero to one and a resolution of 1 mm^2 per pixel. As the smallest expected wavelength was an order of magnitude greater than the mesh size, linear interpolation was deemed acceptable for use. The largest effect to the wavefield data from the linear interpolation was possible flattening of the peaks for the thinnest simulated thickness of 1 mm. An example wavefield image generated in this fashion is shown in Figure 1. Because only external nodes were exported from the ANSYS simulation, no top surface node data was kept where the transducer was located during the simulation. The interpolation of the response to a uniform grid had the effect of smoothing over the circular transducer region, as evident in the bottom left corner of Figure 1. Consequently, any thickness estimations performed in the transducer region are disregarded. It is noted that LDV scans cannot measure the plate surface response beneath the transducer during an experimental test, requiring movement of the transducer to a secondary location if that region is of interest for inspection.

This grayscale image is an example of the input that may be given to the CNN for training or inference. Only the real component of the surface response was used in this study.

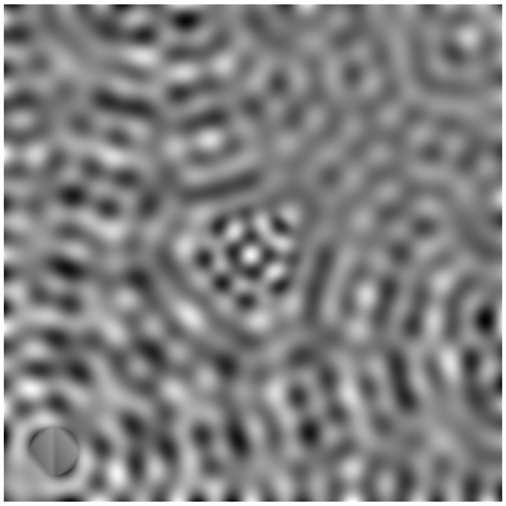


Figure 1: Example of the real-component surface response wavefield image used as input to the CNN

After generating the wavefield image, MATLAB was used to filter and keep all mesh nodes on only the bottom surface of the plate. The local thickness of the plate at each node was inferred from the nodes' Cartesian coordinates and interpolated to the same 400×400 uniform grid as the wavefield image. The local thickness grid was saved as a ground truth label image for the CNN.

All simulation procedures described in this section were connected in a Python automation loop, allowing the programmatic generation of hundreds of labeled wavefield images for training and testing the CNN.

2.3. Training

The Fast.AI (v 1.0.61) and PyTorch (v 1.5.0) machine-learning Python libraries were used in this study for constructing and training the CNN [24], [25]. The ResNet encoder of the CNN was pre-trained on the ImageNet database [26] before being loaded into the wavefield CNN. The same dataset and training procedures as outlined in [13] were used to train the CNN with the Fast.AI library. Namely, the dataset consisted of 503 wavefield images of plates with varying defect sizes, locations, thicknesses, and shapes, as well as varying transducer locations. The dataset was augmented to a size of 2012 images after applying rotations and random white Gaussian noise with a variance of $\sigma^2 = 0.0009$. The amount of noise artificially added to the simulations was representative of noise present in experimental data, e.g. the experimental data in Section 3.3 had an average Gaussian noise variance of $\sigma^2 = 0.0004$. The noise value of the experimental data was calculated by first denoising the experimental dataset using a 2D Wiener filter and a kernel size of 5×5 pixels, then subtracting the denoised image from the experimental image to create an estimated noise image from which the experimental noise variance could be calculated.

This combination of simulation parameters allowed the CNN to learn inherent features within the wavefields rather than the specific shape or size of the defects. Training was performed by minimizing the cross entropy loss function of the model with mini-batch gradient descent. The learning rate optimization tool within Fast.AI was used to smoothly transition the learning rate during training, allowing larger gradient steps at the beginning and more refined steps at the end of training [24].

Due to the programmatic nature of generating a dataset by moving small defects incrementally around the large area of a plate structure, the CNN was trained on significantly more pixels belonging to the nominal plate thickness class (10 mm) than any of the other thickness classes. Table 1 shows the class imbalance in the dataset, given in terms of the percentage of total pixels belonging to each class.

Table 1: Class imbalance in training dataset.

Class (mm)	1	2	3	4	5	6	7	8	9	10
No. Pixels (%)	1.1%	2.0%	3.0%	4.0%	5.0%	4.9%	4.9%	4.9%	4.9%	65.4%

The overarching goal of using the CNN to process ASSESS data is to identify regions in the plate that are smaller than the nominal thickness; however, the nominal plate thickness accounts for 65.4% of the available training data. Future work would account for this imbalance by increasing the presence of non-nominal thicknesses in the training dataset.

It is noted here that the size of the images the CNN was trained on aligned exactly with the size of all plates in the simulated dataset, namely plates of size 400 mm × 400 mm, with a resolution of 1 mm per pixel. A side effect of this approach is that the CNN was only exposed to the free boundary conditions of the simulated plates along the edges of the image during training. Future work would expose the CNN to a greater variety of real-world boundary conditions and scenarios in the training set.

2.4. Simulation test set

This study seeks to evaluate the performance of the CNN in characterizing wavefields from ASSESS data. Of particular interest is the performance of the CNN with respect to detecting defects of varying sizes and material thickness. A collection of 10 mm thick, 400 mm × 400 mm aluminum alloy plates was programmatically generated in Solidworks, where each plate contained one thinning defect near the center of the plate. The harmonic responses of the plates were gathered according to the technique described in section 2.2. The local percent reduction in plate thickness at the defect varied from 10% to 90% in 10%

increments. The characteristic length (i.e. the largest dimension) of the defect varied from 10 mm to 100 mm in 10 mm increments, for a total test set size of 90. Figure 2 shows an example of the test set geometry.

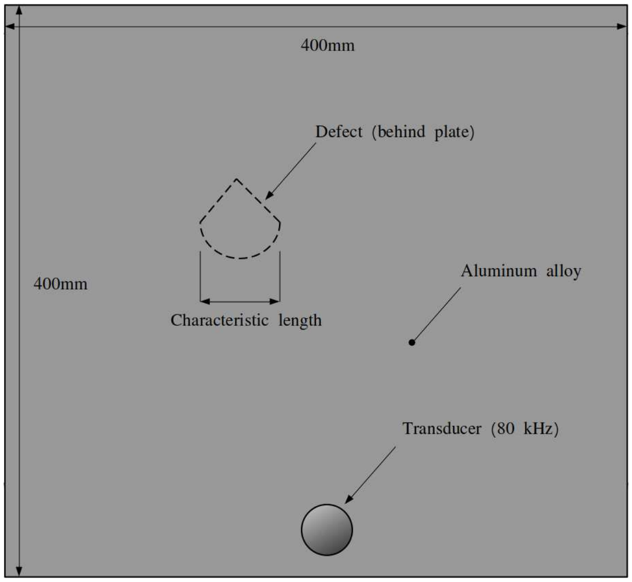


Figure 2: Test set plate geometry

It is noted that the dataset used to train the CNN did not contain a plate with the same defect shape or location as in the test set. The training dataset also did not contain a plate with the transducer in the same location as shown in Figure 2. The training dataset did contain plates with no defects, so the testset will exclude those cases here. All parameters other than defect thickness and size were kept constant in the test set. All simulations in the fully independent test set contained the same boundary conditions along the edges of the plate as was used in the training dataset.

2.5. Performance metrics

The intersection over union (IoU) value is the most commonly reported metric when evaluating the performance of image segmentation models. The IoU value is a measure of how much overlap there is between a predicted image and the ground truth image, with higher values tending to more accurate prediction results. For binary images, the IoU value is defined as:

$$IoU = \frac{TP}{TP + FP + FN} \tag{4}$$

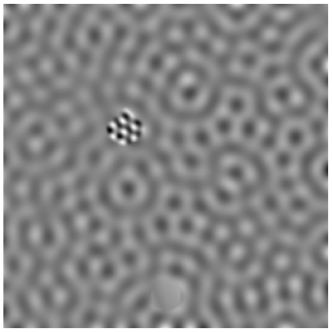
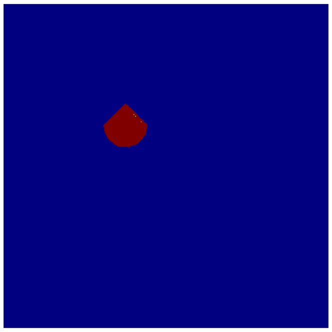
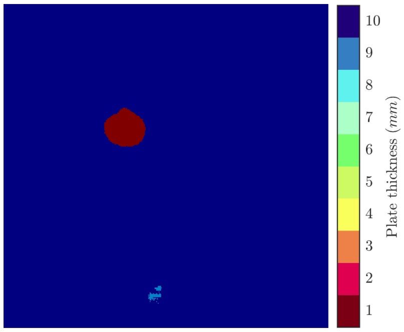
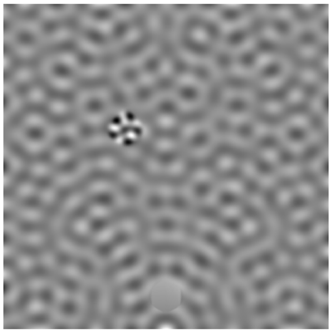
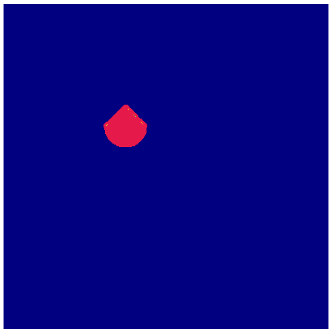
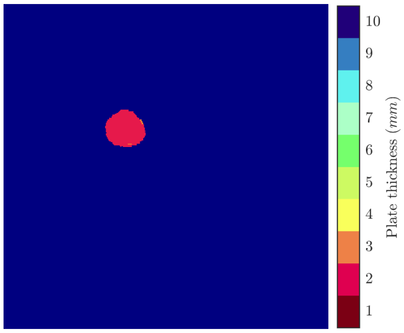
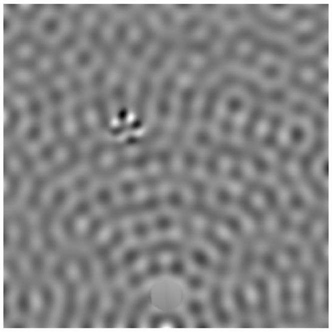
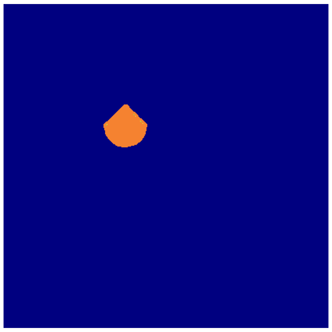
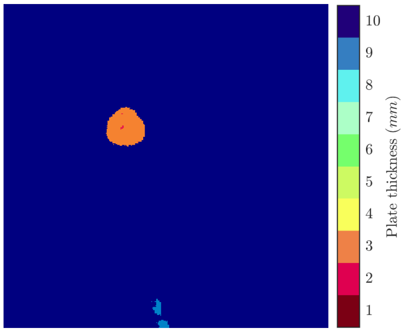
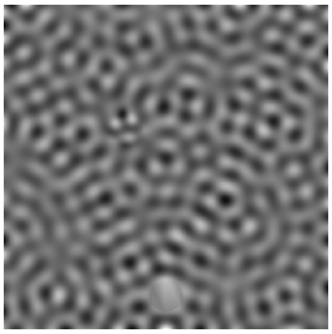
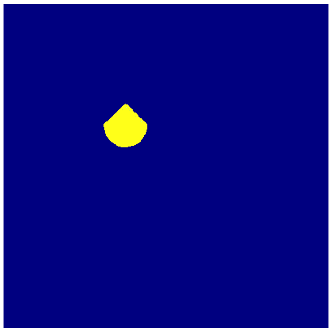
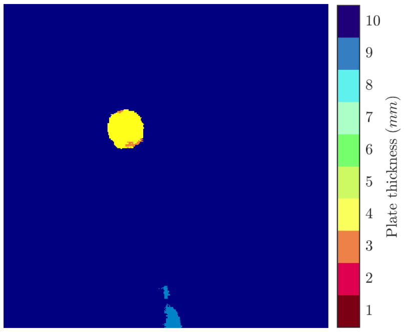
where TP is true positive, FP is false positive, and FN is false negative. IoU provides a more reasonable metric of model accuracy compared to raw pixel accuracy, since false positives penalize the metric. An IoU of 1 corresponds to a perfect prediction. For non-binary images with multiple class values possible for each pixel, the IoU of each class is calculated with (4), and then averaged across all classes. For the purposes of this study, the performance of the CNN in predicting the location and thickness of the defect will be evaluated using only the IoU values of the classes present in the defect regions in every test image.

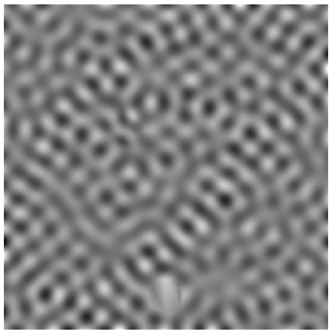
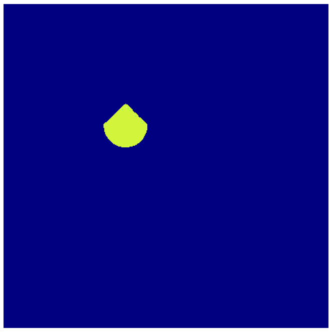
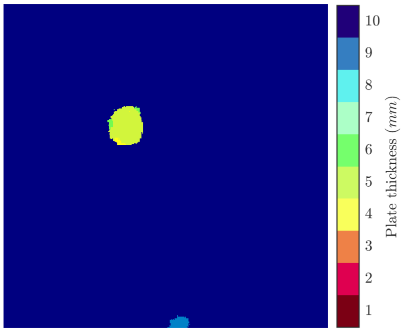
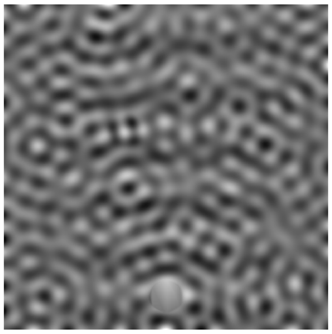
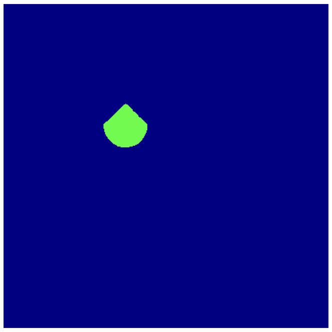
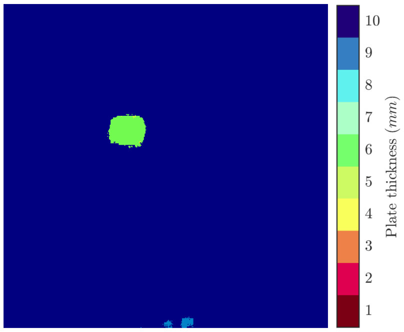
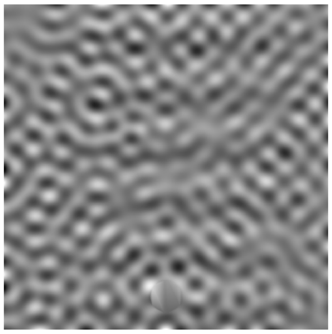
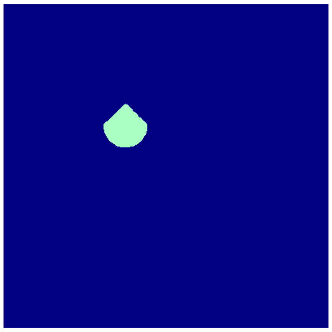
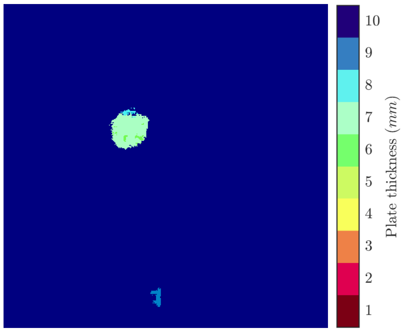
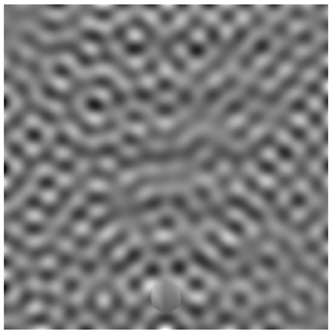
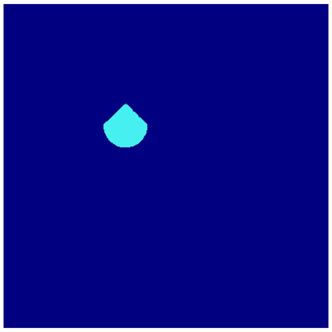
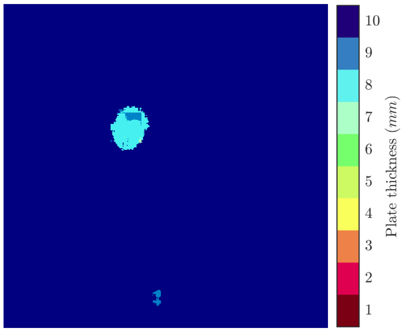
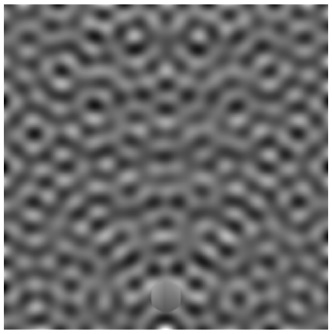
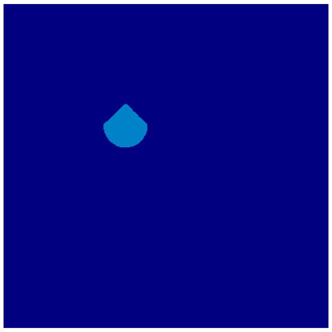
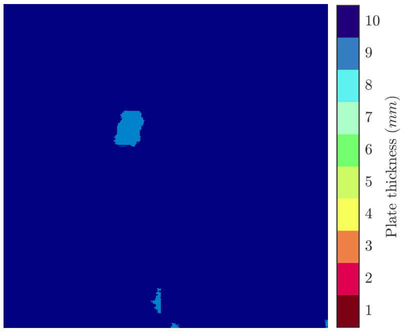
3. Results

3.1. Simulation test set

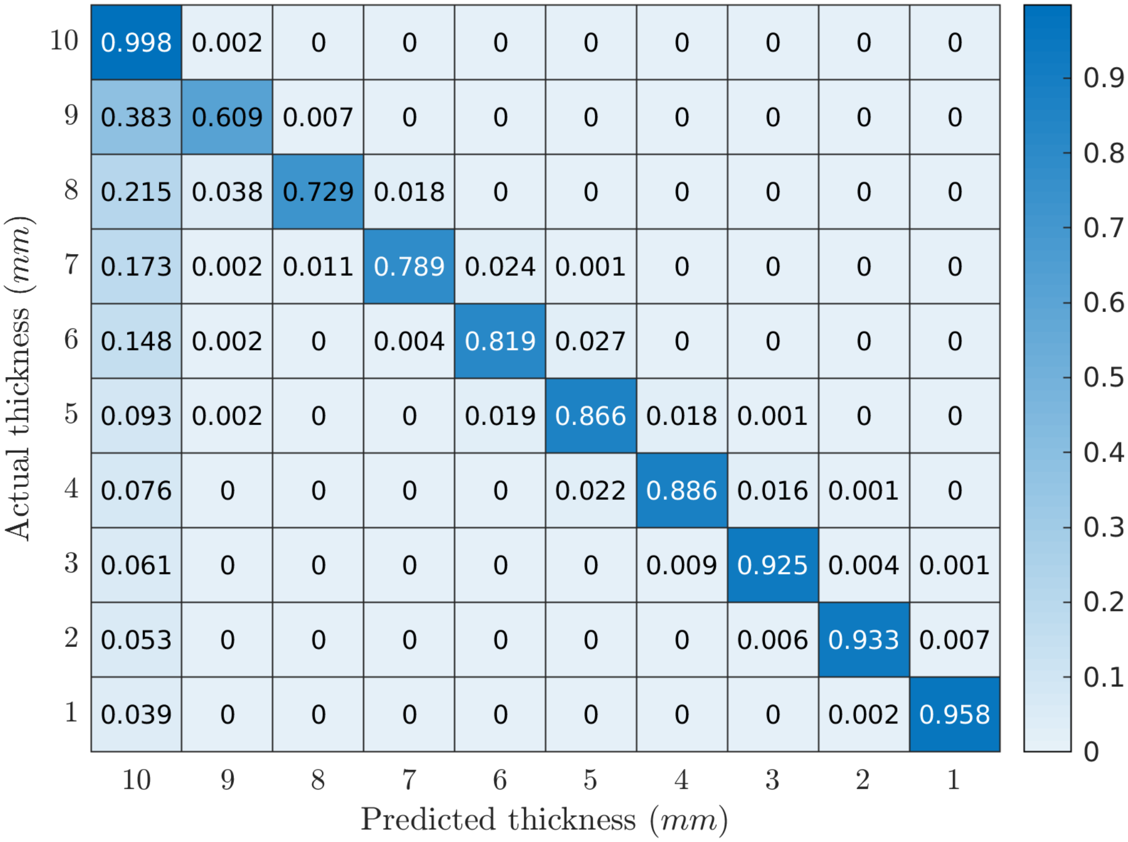
This section presents the CNN thickness prediction results from the simulated test set described in Section 2.4. In each row of Table 2, the wavefield image for each test case is shown along with the ground truth segmentation image and the CNN thickness prediction image. The left-most column identifies the test case by plate thickness at the defect t (mm) and defect characteristic length L (mm). The results are shown here for a defect of size 50 mm for all thickness classes. The IoU value for each case is included. The full table of all CNN prediction results is provided in Appendix A.

Table 2: CNN thickness prediction results on the test set (50 mm only)

t	L	IoU	Wavefield	Ground Truth	Prediction
1	50	90%			
2	50	88%			
3	50	88%			
4	50	80%			

t	L	IoU	Wavefield	Ground Truth	Prediction
5	50	73%			
6	50	72%			
7	50	66%			
8	50	63%			
9	50	51%			

Despite the high class imbalance in the training and test sets, the CNN still shows reasonable performance in identifying non-nominal plate thickness regions. The normalized confusion matrix over the test set is shown in Figure 3.



Actual thickness (mm) \ Predicted thickness (mm)	10	9	8	7	6	5	4	3	2	1
10	0.998	0.002	0	0	0	0	0	0	0	0
9	0.383	0.609	0.007	0	0	0	0	0	0	0
8	0.215	0.038	0.729	0.018	0	0	0	0	0	0
7	0.173	0.002	0.011	0.789	0.024	0.001	0	0	0	0
6	0.148	0.002	0	0.004	0.819	0.027	0	0	0	0
5	0.093	0.002	0	0	0.019	0.866	0.018	0.001	0	0
4	0.076	0	0	0	0	0.022	0.886	0.016	0.001	0
3	0.061	0	0	0	0	0	0.009	0.925	0.004	0.001
2	0.053	0	0	0	0	0	0	0.006	0.933	0.007
1	0.039	0	0	0	0	0	0	0	0.002	0.958

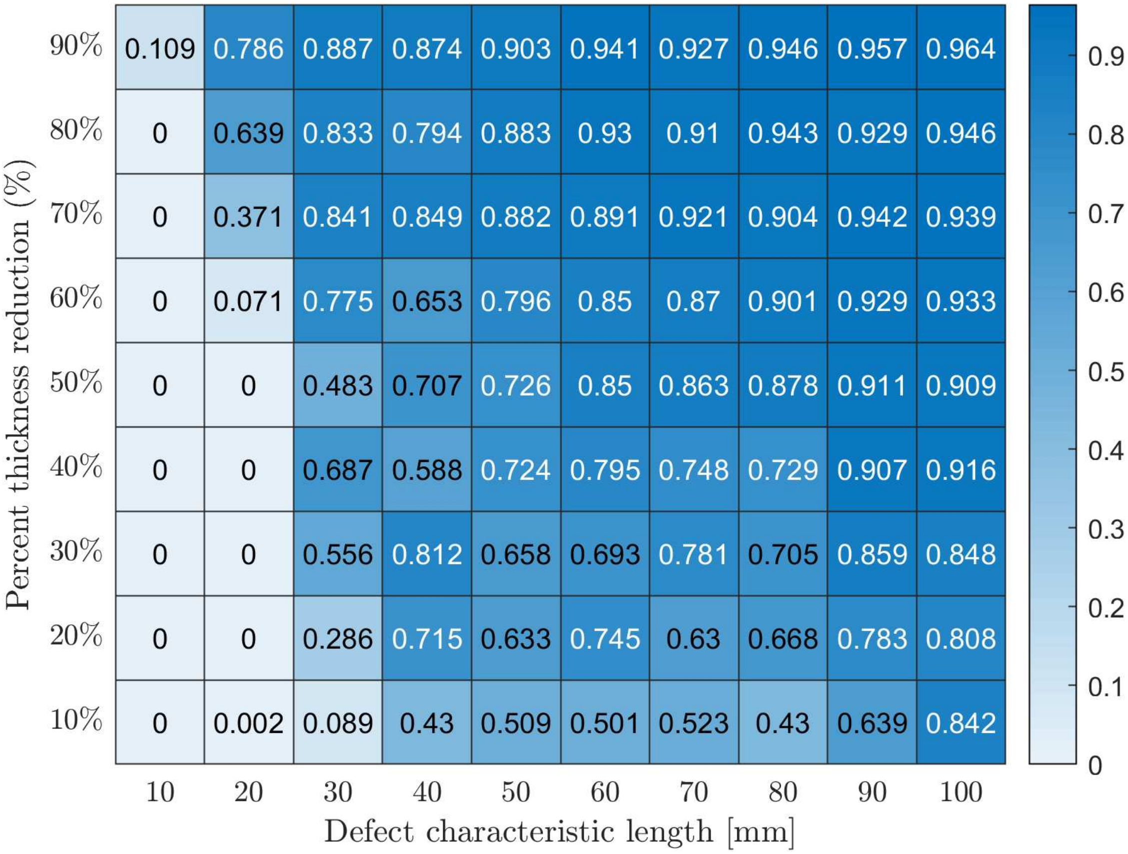
Figure 3: Normalized confusion matrix over test set

The percentage of pixels predicted correctly for a given class are shown along the blue diagonal. All percentages add up to 100% across the rows of the confusion matrix. When the CNN predicted a given pixel incorrectly, it most commonly predicted it as belonging to the plate thickness class (10 mm) or to one of the directly adjacent thickness classes.

Table 2 and the full results table in Appendix A characterize the performance of the CNN with respect to thickness reduction and defect size. When the reduction in plate thickness at the defect is high, the wavefield correspondingly exhibits sharper changes in wavenumber at the defect region. The high difference in wavenumber between the defect region and the rest of the plate results in the defect being visually distinctive within the wavefield, and the CNN predicts the boundaries of the defect more accurately. As the reduction in plate thickness decreases, the wavenumber in the defect region approaches the nominal plate wavenumber and it becomes less visually distinct, tending to less accurate CNN predictions. The CNN also shows increasing accuracy at a given defect thickness as the size of the defect increases. The accuracy of the CNN in predicting defect thickness and location is evaluated by the IoU metric in (4) using only the pixels belonging to the defect's thickness class. The defect IoU for each of the test cases is presented in the chart in Figure 4. It is noted that the plate thickness at each defect is reported here as a percent thickness reduction, given by

$$\%\ \text{thickness reduction} = \frac{t_n - t_d}{t_n} \times 100\% \tag{5}$$

for a nominal plate thickness t_n and a plate thickness at the defect of t_d.



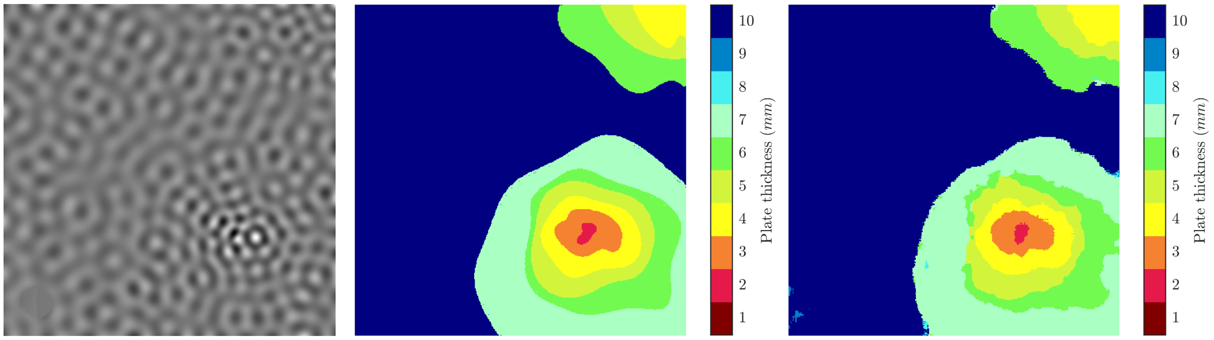
Percent thickness reduction (%)	10	20	30	40	50	60	70	80	90	100
90%	0.109	0.786	0.887	0.874	0.903	0.941	0.927	0.946	0.957	0.964
80%	0	0.639	0.833	0.794	0.883	0.93	0.91	0.943	0.929	0.946
70%	0	0.371	0.841	0.849	0.882	0.891	0.921	0.904	0.942	0.939
60%	0	0.071	0.775	0.653	0.796	0.85	0.87	0.901	0.929	0.933
50%	0	0	0.483	0.707	0.726	0.85	0.863	0.878	0.911	0.909
40%	0	0	0.687	0.588	0.724	0.795	0.748	0.729	0.907	0.916
30%	0	0	0.556	0.812	0.658	0.693	0.781	0.705	0.859	0.848
20%	0	0	0.286	0.715	0.633	0.745	0.63	0.668	0.783	0.808
10%	0	0.002	0.089	0.43	0.509	0.501	0.523	0.43	0.639	0.842

Figure 4: Defect IoU for every test case. IoU ranges from 0 to 1.

The nominal wavelength in the 10 mm thickness plate was $O(10\,\text{mm})$. As a result, defects with a 10 mm characteristic length go almost completely undetected by the CNN, while 100 mm defects show IoUs over 80%. Defects at 90% reduction are the easiest to detect, with IoUs over 80% across most defect sizes. Defects at 10% reduction show the worst performance at around 50% IoU for most defect sizes. As the wavefield measurements are dependent on both the local material thickness and the excitation frequency, it is possible that higher excitation frequencies could improve these results, but such investigation was outside the scope of this study.

3.2. Additional simulation test cases

The previous test set showed a parameter study for a relatively simple defect geometry of varying thicknesses and sizes. This section highlights the generalizability of the CNN-based thickness estimation technique to more complex geometries. A plate was modeled with multiple nested defects along the edges of the plate and simulated according to the procedure in Section 2.2. One large defect in the bottom right corner of the plate had plate thicknesses ranging from 7 mm incrementally down to 2 mm near its center. The arbitrary shape and distribution of thicknesses in the defect might be more representative of a real-world corrosion or indentation in the back of the plate. A similar “corrosion” defect was applied to the top right corner of the plate, with thicknesses ranging incrementally from 6 mm to 4 mm. The wavefield image, along with the ground truth and CNN thickness prediction maps, are shown in Figure 5.



(a) Wavefield **(b)** Ground Truth **(c)** Prediction

Figure 5: CNN performance on a complex geometry

For a test case with multiple thickness classes present in the image, the accuracy metric is reported by calculating the IoU of each thickness class individually (Eq. (4)), and averaging over the number of thickness classes present in the defect regions:

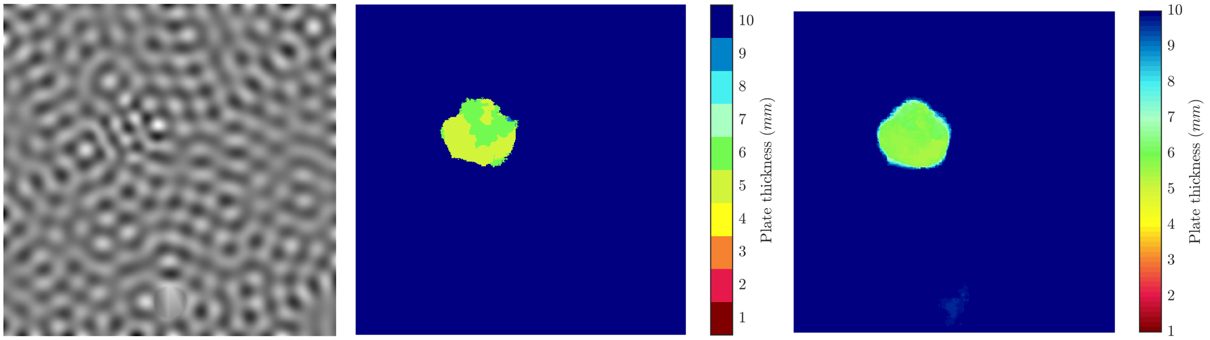
$$IOU_{avg} = \frac{1}{n}\sum_{i=1}^{n} IOU_i \tag{6}$$

where IOU_i is the accuracy metric of the i^{th} thickness class for n total thickness classes under consideration. For this test case, the defect regions are composed of 6 classes, namely $2-7$ mm. So $n=6$ and IOU_i is calculated at each incremental thickness class i from 2 to 7 mm. The average IoU for this test case is 79.9%. The CNN demonstrates the ability to accurately predict multiple thicknesses in complex, nested regions, as well as along the boundaries of the plate.

Next, a simple geometry, as described in Section 2.4, was created with a characteristic length of 90 mm and a defect thickness of 5.5 mm. The CNN thickness prediction map was generated using only the real component of the wavefield. Due to the limitations of discrete classification, the CNN predicts the 5.5 mm defect thickness as primarily 5 mm with several patches of 6 mm. In order to resolve the sub-millimeter thickness and to provide the CNN with information of the full wavefield, the following procedure was used:

1. The real-component wavefield image was treated as a single snapshot in the harmonic cycle, $\theta \in [0, 2\pi]$. We treat this snapshot as the beginning of the cycle at $\theta = 0°$.
2. A new snapshot of the harmonic cycle was generated by incrementing θ by 10°. This new snapshot of the wavefield velocity response follows from Eq. (3) at the new θ location.
3. A total of 36 wavefield images was gathered in this fashion, corresponding to snapshots at 10° increments across the 360° complex plane.
4. Each of the 36 images was fed through the CNN separately to generate 36 unique thickness predictions of the same plate. Each prediction image contained discrete integer class values at each pixel, as described in section 2.
5. A simple arithmetic mean was taken across all 36 prediction images to obtain a single, averaged thickness prediction map.

The single and phase-shift averaged CNN thickness predictions are shown in Figure 6.

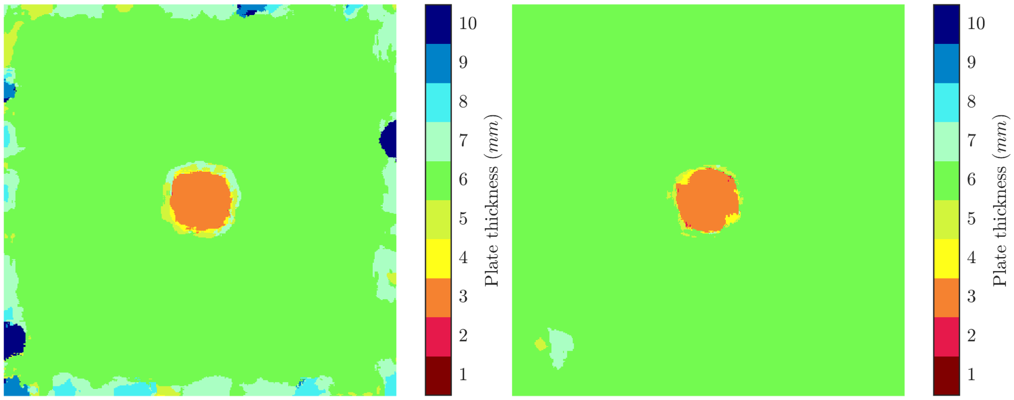


(a) Wavefield **(b)** Single prediction **(c)** Phase-shifted average prediction

Figure 6: CNN performance on a sub-millimeter defect

The average plate thickness prediction in the defect region of the phase-shift average image was 5.7 mm, which is a 3.6% difference from the true value. The average thickness for the single prediction was 5.4 mm, (1.8% difference). Even though the CNN is limited to discrete classification, it shows reasonable performance in locating and characterizing defects that do not align on integer thickness values. In addition, sub-millimeter resolution can be approximated by averaging CNN inference over the full 360 degree complex wavefield.

Insofar, all simulation test cases have been performed on 400 mm × 400 mm plates with free boundary conditions that align along the edges of the wavefield image, as this was the condition of all images in the CNN's training dataset. A 0.91 m × 0.91 m aluminum plate with a 6 mm thickness and a 3 mm thick circular defect at the center of the plate was modeled and simulated according to the procedure in Section 2.2. The simulation results were cropped to a central 400 mm × 400 mm region, which included the circular defect. Another simulation was performed in exactly the same way, except the size of the plate was 400 mm × 400 mm, and no cropping was needed. Both simulations used free boundary conditions along the edges of the plate. The CNN thickness prediction maps for both cases are shown in Figure 7.



(a) 0.91 m × 0.91 m (cropped) **(b)** 400 mm × 400 mm (no crop)

Figure 7: CNN thickness prediction maps for two 6 mm plates of different sizes

The CNN accurately predicts the nominal plate thickness and the defect region in both cases; however, the prediction on the cropped region of the larger plate shows false predictions along the edges of the image. The only difference between these two cases is the presence or absence of a physical plate boundary along the edges of the image. The most likely cause of the false edge predictions in the cropped image is the absence of the boundary effects seen in the un-cropped image, as well as those seen in the entire training dataset. The CNN was only trained to expect physical plate boundaries along the edges of the image. When presented with a cropped region of a much larger plate, the CNN loses generalizability.

Despite the boundary condition limitation in the training dataset used in this study, the CNN has shown the ability to generalize learned features in simulated wavefield measurements to a variety of defect sizes, thicknesses, and geometries. For test cases with the same boundary conditions as learned by the CNN during training, the CNN has shown reasonable performance in predicting thickness in regions normally limited by complex or unknown Lamb wave dispersion relationships, such as along the edges of the scan region.

3.3. Experimental test set

This section will present the application of the CNN to an experimental test set to evaluate the generalizability of the simulation-trained network to ASSESS experimental data. The test specimen consisted of a large aluminum plate (6.35 mm thick, 1.2 m × 1.2 m) with two elongated defect regions milled out on the back surface with a 2.5 cm end-mill. The first defect, referred to as "deep cut", was created by removing amounts of material ranging from 2.3 mm to 4.2 mm. The second defect, referred to as "shallow cut", was created in the same way with material removal ranging from 0.75 mm to 1.2 mm. The elongated shapes were approximately 2.5 cm wide and 20.3 cm long. The aluminum test specimen is shown in Figure 8.

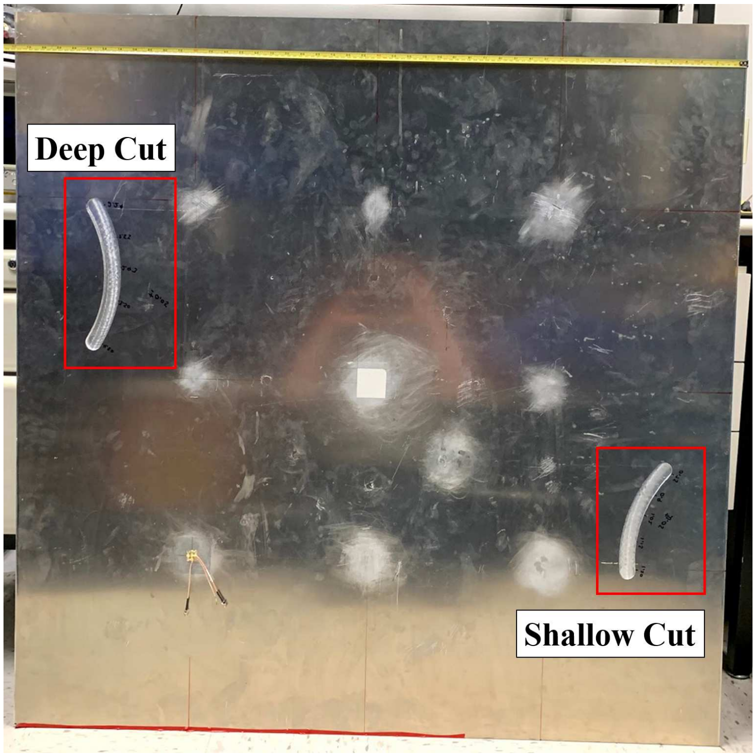


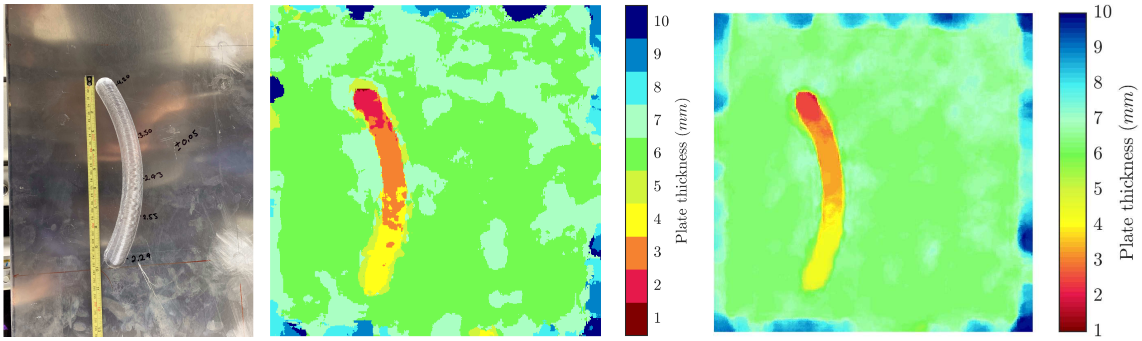
Figure 8: Aluminum test specimen with two milled out defects

Experimental data was collected using an ASSESS scanning system, which uses two steering mirrors and an LDV to record the steady-state response of a structure excited by an ultrasonic transducer [3]. The ASSESS system recorded the steady-state velocity response of the top surface of the plate. The ultrasonic transducer was clamped to the corner of the plate and excited at 80 kHz, which was the same excitation frequency as all the simulations used to train the neural network. The data was collected in two regions – one surrounding each defect – spanning across at least a 400 mm × 400 mm region with a 1 mm^2 pixel resolution.

Independent analysis of the collected data was performed using the original ASSESS wavenumber estimation algorithm [7, 8]. This algorithm passes the matrix of ASSESS results through narrow-band wavenumber filters. Then, the spatial envelope of each wavenumber filter output is calculated and the wavenumber corresponding to the spatial envelope with the maximum amplitude at each point is selected as the estimated wavenumber. Material-dependent dispersion curves were used to map wavenumber to thickness. In the thickness and frequency ranges of interest, the specimen's response was dominated by the zero-order antisymmetric (A0) mode. Mapping between wavenumber and thickness was accomplished through interpolation along the A0 mode.

This wavenumber estimation tool was used to quantitatively evaluate the data for any environmental discrepancies that are inherent to physical testing. These variabilities can include material property variation within a structure, a spatial sample rate that slightly deviates from 1 mm^2, errors in physical distance and look angle of the ASSESS system, and other component-level discrepancies. A system-wide calibration factor was derived by spatially resampling the experimental data to align the calculated wavenumber to the healthy specimen's expected wavenumber through interpolation on the A0 mode. This calibration factor compensates for any of the measurement discrepancies listed above. The deep cut dataset required a scale factor of 0.985 and the shallow cut required a scale factor of 0.963. The need for unique calibration factors may in part be caused by the fact that the ASSESS system was repositioned between measurements.

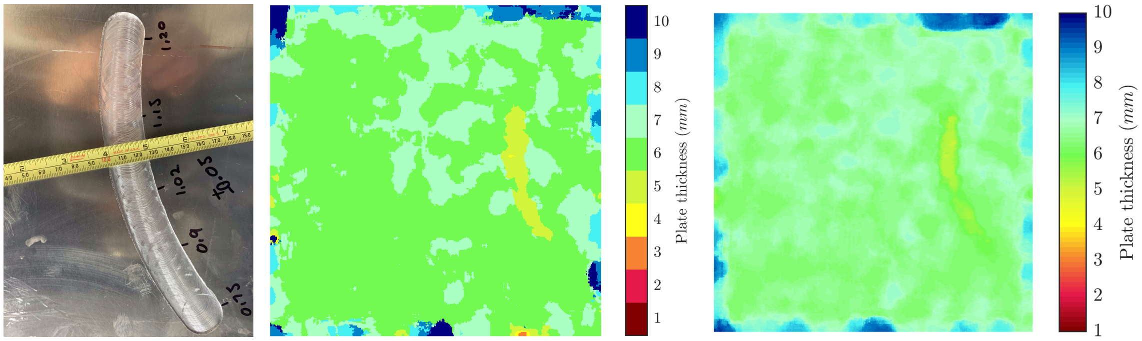
The CNN thickness prediction map was generated for both defects using just the real component of the wavefield, as well as using a phase-shift average, as was described for the test case in Figure 6. Figure 9 shows the thickness prediction maps on the deep cut defect. The markings next to the defect in Figure 9a indicate the amount of material removed at that location (in mm). The defect region shows a plate thickness of 4 mm that gradually worsens to 2 mm at the tip of the defect, corresponding to the 2.3 mm of material removed at the bottom of the defect and the 4.2 mm of material removed at the top of the defect. Several false predictions can be seen around the edges of the 400 mm × 400 mm cropped region, similarly to the simulation results presented in Section 3.2. The average nominal plate thickness prediction in the healthy region of the phase-shift average image was 6.45 mm, which is a 1.6% difference from the true value. The average thickness in the healthy region of the single prediction image was 6.22 mm, (2.0% difference).



(a) Ground truth image **(b)** Single CNN thickness prediction **(c)** Phase-shift averaged prediction

Figure 9: CNN thickness prediction map for deep cut defect

Figure 10 shows the CNN thickness prediction maps for the shallow cut defect.



(a) Ground truth image **(b)** Single CNN thickness prediction **(c)** Phase-shift averaged prediction

Figure 10: CNN thickness prediction map for shallow cut defect

The defect region is identified where the nominal plate thickness was reduced by 1 mm to a thickness of approximately 5 mm. The average nominal plate thickness prediction in the healthy region of the phase-shift average image was 6.55 mm, which is a 3.1% difference from the true value. The average thickness in the healthy region of the single prediction image was 6.28 mm, (1.1% difference). This test case shows the same limitations in predicting the edges of the interrogation region.

Despite the boundary condition limitations described previously, the CNN has shown reasonable performance in locating and characterizing defects from experimental ASSESS data using local plate thickness estimation.

4. Conclusions

This paper has presented a CNN model for semantic image segmentation trained to estimate plate thickness from full-field, steady-state ultrasonic wavefield images. ASSESS data was simulated in FEA software and processed by the CNN to produce plate thickness segmentation maps for defect localization and characterization. The trained CNN has demonstrated accurate thickness prediction without reliance on the 2D FFT or Lamb wave dispersion relationships, and has shown a robust generalizability to experimental ASSESS data despite a fully-simulated training dataset. The performance limits of the CNN were explored with respect to defect size and thickness reduction. The CNN showed greater accuracy for larger defects with greater thickness reduction, and declining accuracy for smaller defects with smaller thickness reductions. A key limitation in the CNN is its confinement to discrete classification. Future work is needed to predict continuous thickness maps with greater class balance in the training dataset. Further studies should also expand the training dataset to account for more real-world scenarios, including more complex plate geometries and boundary conditions.

Acknowledgements

Research presented in this article was supported by the Laboratory Directed Research and Development Program of Los Alamos National Laboratory (LANL) under project number 20190580ECR and the Engineering Institute's Los Alamos Dynamics Summer School. The Engineering Institute is a research and education collaboration between LANL and the University of California San Diego's Jacobs School of Engineering. This collaboration seeks to promote multidisciplinary engineering research that develops and integrates advanced predictive modeling, novel sensing systems, and new developments in information technology to address LANL mission relevant problems. LANL is operated by Triad National Security, LLC (Contract No. 89233218CNA000001) for the National Nuclear Security Administration of the U.S. Department of Energy.

References

[1] J. E. Michaels, Ultrasonic wavefield imaging: Research tool or emerging nde method?, AIP Conference Proceedings 1806 (1) (2017) 020001. `doi:10.1063/1.4974542`.

[2] M. Ruzzene, Frequency–wavenumber domain filtering for improved damage visualization, Smart Materials and Structures 16 (6) (2007) 2116–2129.

[3] E. B. Flynn, G. S. Jarmer, High-speed, non-contact, baseline-free imaging of hiddden defects using scanning laser measurements of steady-state ultrasonic vibration, in: Structural Health Monitoring 2013: Proc. of the Ninth Int. Workshop on Structural Health Monitoring, Vol. 1, 2013, pp. 1186–1193.

[4] M. D. Rogge, C. A. Leckey, Characterization of impact damage in composite laminates using guided wavefield imaging and local wavenumber domain analysis, Ultrasonics 53 (7) (2013) 1217–1226.

[5] P. Huthwaite, F. Simonetti, High-resolution guided wave tomography, Wave Motion 50 (5) (2013) 979 – 993. `doi:https://doi.org/10.1016/j.wavemoti.2013.04.004`.

[6] P. Aryan, A. Kotousov, C. T. Ng, B. S. Cazzolato, A baseline-free and non-contact method for detection and imaging of structural damage using 3D laser vibrometry, Structural Control and Health Monitoring 24 (4) (Jun. 2017).

[7] E. B. Flynn, S. Y. Chong, G. J. Jarmer, J.-R. Lee, Structural imaging through local wavenumber estimation of guided waves, NDT & E International 59 (2013) 1–10. `doi:10.1016/j.ndteint.2013.04.003`.

[8] E. B. Flynn, Embedded multi-tone ultrasonic excitation and continuous-scanning laser doppler vibrometry for rapid and remote imaging of structural defects, in: EWSHM - 7th European Workshop on Structural Health Monitoring, Nantes, France, 2014.
URL `https://hal.inria.fr/hal-01021054`

[9] E. B. Flynn, N. D. Stull, Toward utilizing full-field laser-ultrasound for practical nondestructive inspection with acoustic wavenumber spectroscopy (2018). `doi:10.1109/ultsym.2018.8579833`.

[10] N. M. O'Dowd, D. Han, L. Kang, E. B. Flynn, Exploring the performance limits of full-field acoustic wavenumber spectroscopy techniques for damage detection through numerical simulation, in: 8th European Workshop on Structural Health Monitoring, 2016.

[11] O. Mesnil, H. Yan, M. Ruzzene, K. Paynabar, J. Shi, Fast wavenumber measurement for accurate and automatic location and quantification of defect in composite, Structural Health Monitoring 15 (2) (2016) 223–234.

[12] N. Stull, E. Flynn, M. Mascarenas, On the theoretical limitations in estimating thickness of a plate-like structure from a full-field single-tone response lamb wave measurement, Ultrasonics 108 (2020) 106230. `doi:10.1016/j.ultras.2020.106230`.

[13] J. D. Eckels, I. F. Fernandez, K. Ho, N. Dervilis, E. M. Jacobson, A. J. Wachtor, Application of a U-net convolutional neural network to ultrasonic wavefield measurements for defect characterization, in: Rotating Machinery, Optical Methods & Scanning LDV Methods, Volume 6, Springer International Publishing, 2022, pp. 167–181. `doi:10.1007/978-3-030-76335-0_18`.

[14] I. Virkkunen, T. Koskinen, O. Jessen-Juhler, J. Rinta-Aho, Augmented ultrasonic data for machine learning, arXiv preprint arXiv:1903.11399 (2019).

[15] Q. Li, Y. Da, Y. Zhang, B. Wang, D. Liu, Z. Qian, A novel combination of theoretical analysis and data-driven method for reconstruction of structural defects (2020). `arXiv:2009.06276`.

[16] M. Meng, Y. J. Chua, E. Wouterson, C. P. K. Ong, Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks, Neurocomputing 257 (2017) 128 – 135.

[17] A. A. Ijjeh, S. Ullah, P. Kudela, Full wavefield processing by using fcn for delamination detection, Mechanical Systems and Signal Processing 153 (2021) 107537. `doi:https://doi.org/10.1016/j.ymssp.2020.107537`.

[18] H. Song, Y. Yang, Noncontact super-resolution guided wave array imaging of subwavelength defects using a multiscale deep learning approach, Structural Health Monitoring (2020). `doi:10.1177/1475921720942958`.

[19] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, M. Pietikäinen, Deep learning for generic object detection: A survey, International journal of computer vision 128 (2) (2020) 261–318.

[20] F. Lateef, Y. Ruichek, Survey on semantic segmentation using deep learning techniques, Neurocomputing 338 (2019) 321–348.

[21] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, Pattern Recognition 77 (2018) 354 – 377.

[22] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241.

[23] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

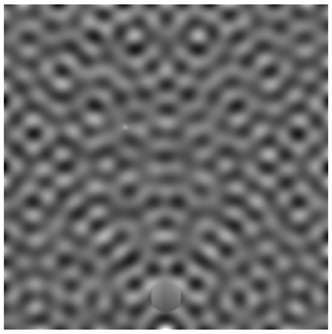
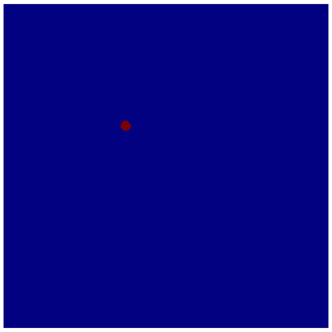
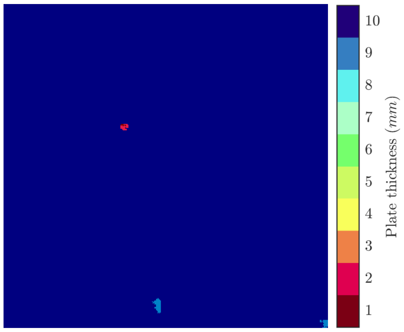
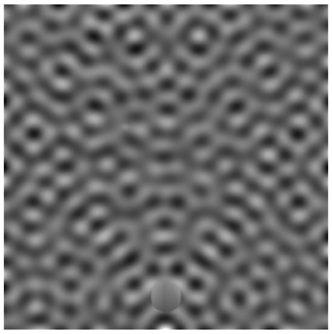
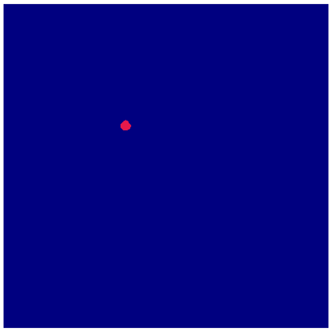
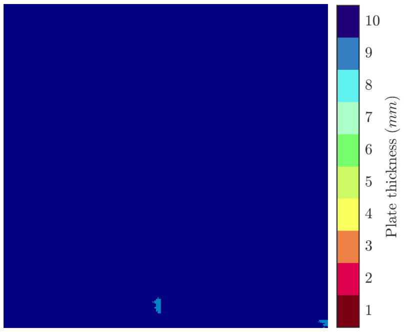
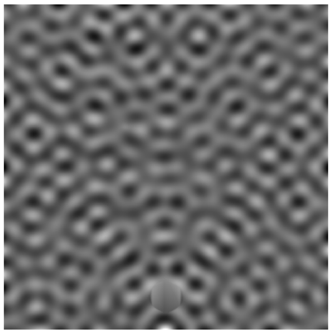
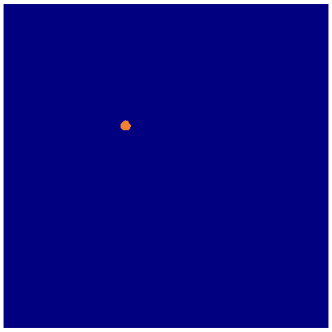
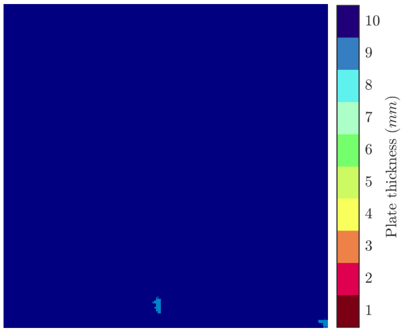
[24] J. Howard, S. Gugger, Fastai: A layered API for deep learning, Information 11 (2) (2020) 108.

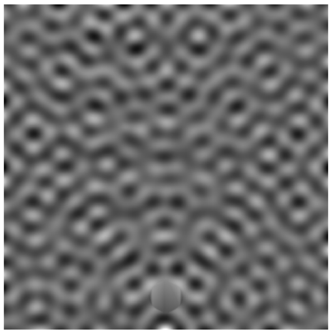
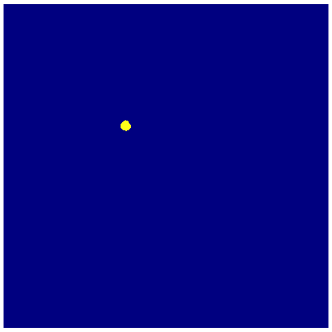
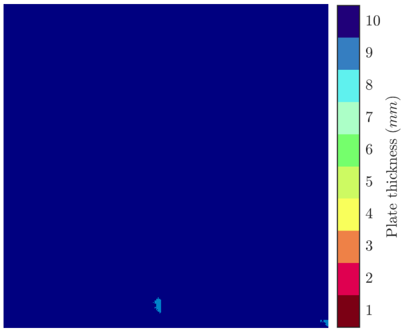
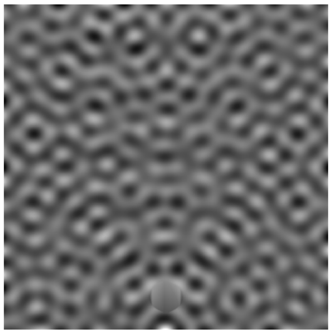
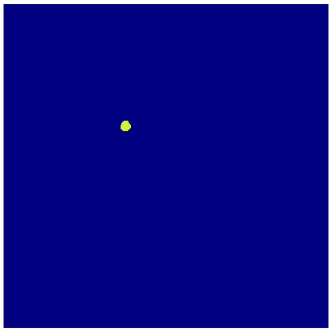
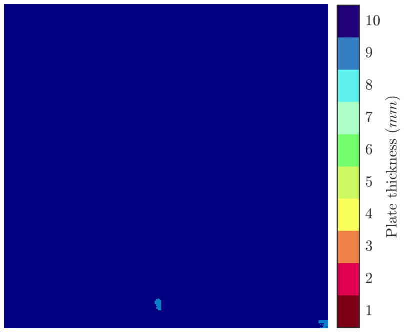
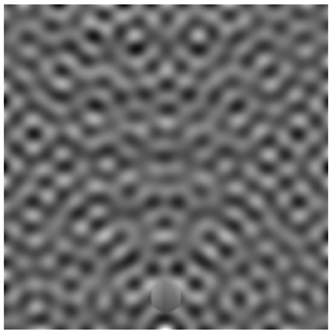
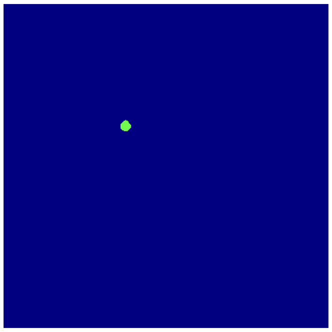
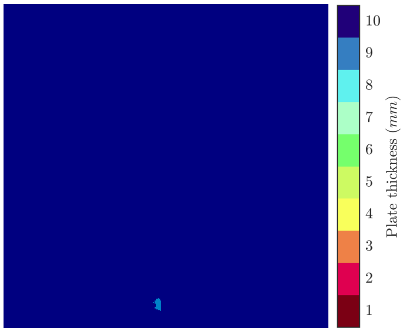
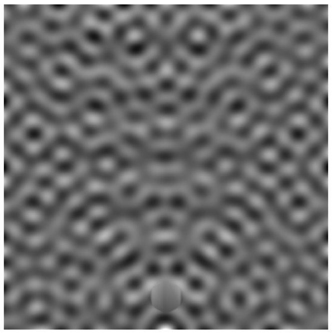
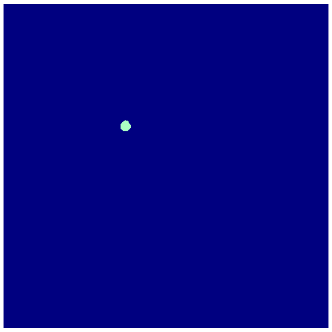
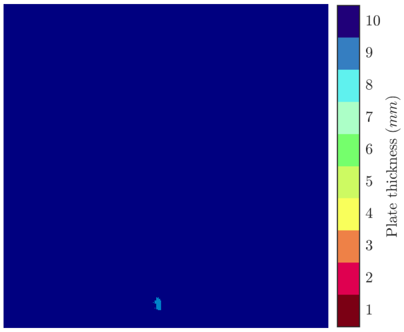
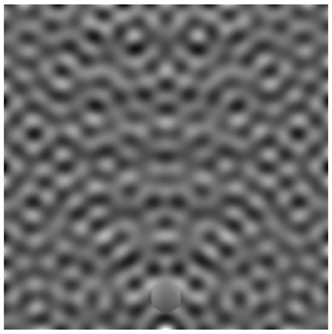
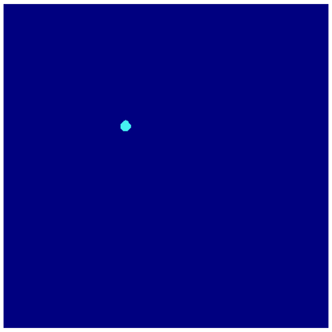
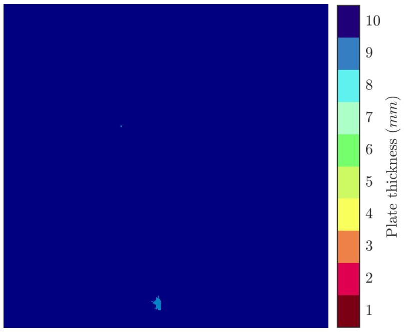
[25] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An imperative style, high-performance deep learning library, in: Advances in Neural Information Processing Systems 32, 2019, pp. 8024–8035.

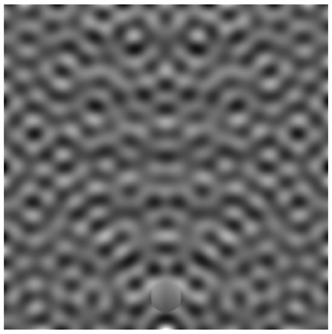
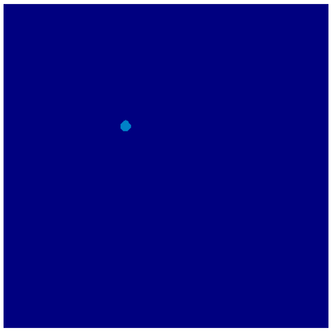
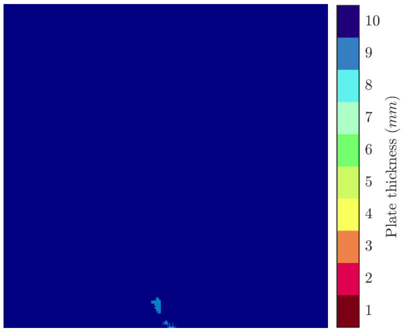
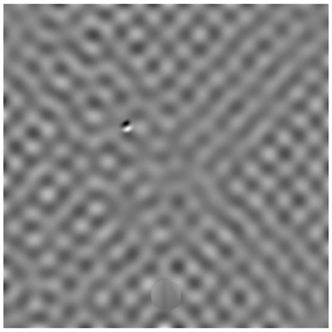
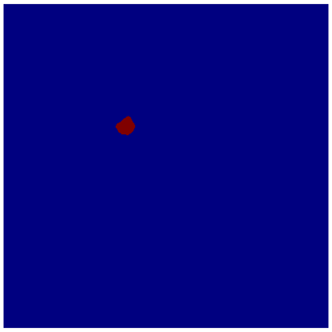
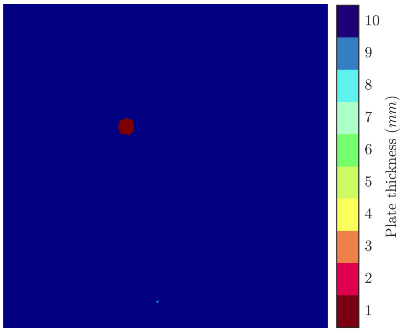
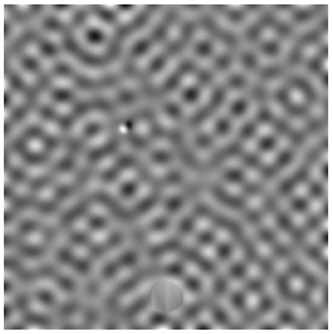
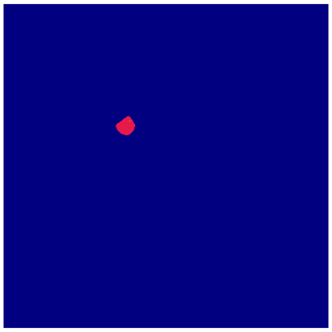
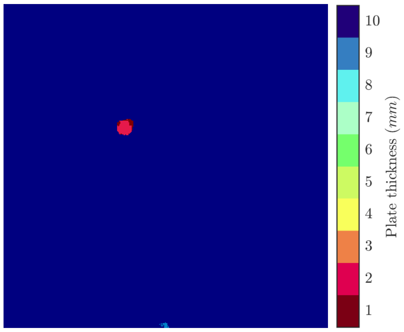
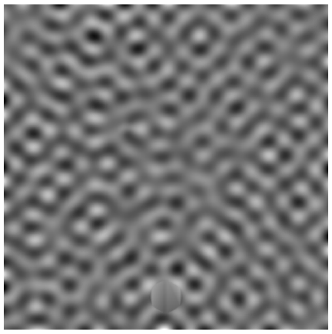
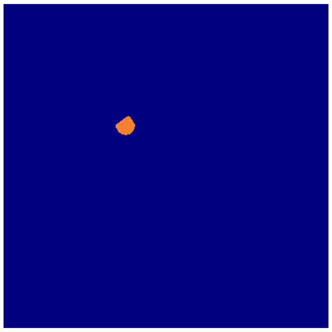
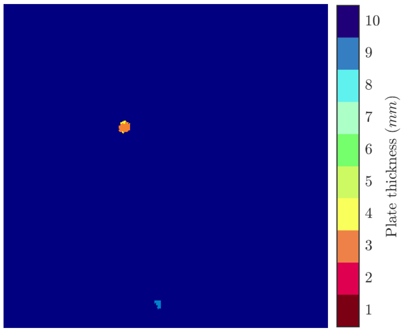
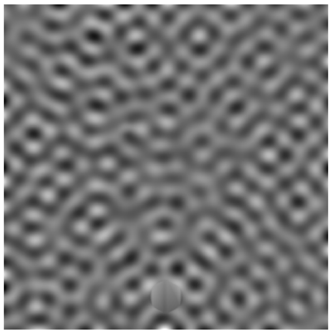
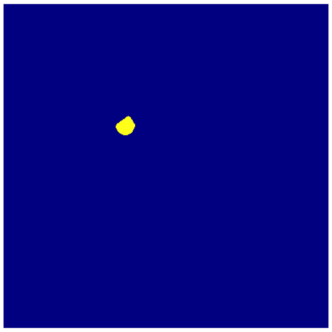
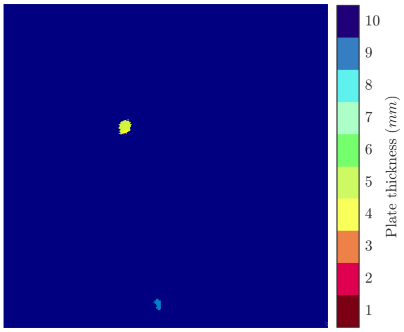
[26] J. Deng, W. Dong, R. Socher, L. jia Li, K. Li, L. Fei-fei, ImageNet: A large-scale hierarchical image database, in: In CVPR, 2009.

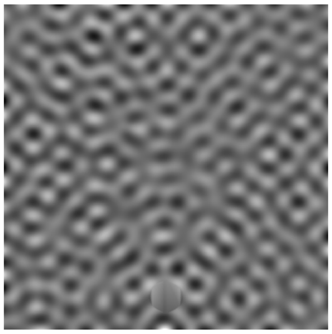
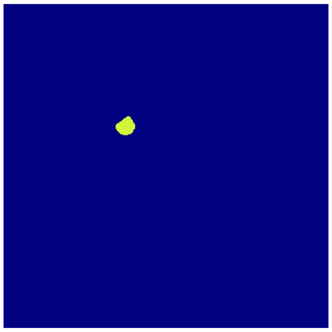
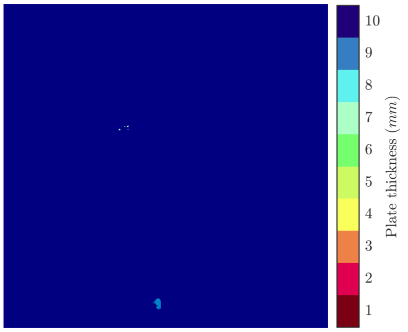
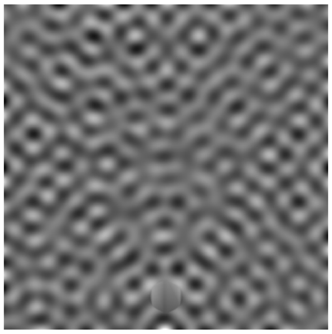
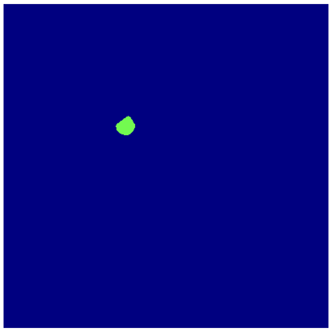
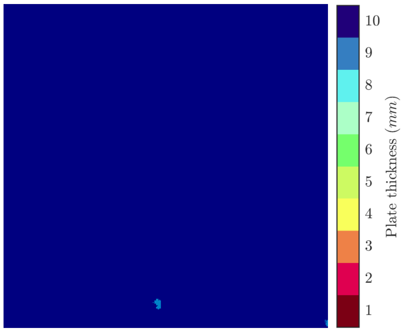
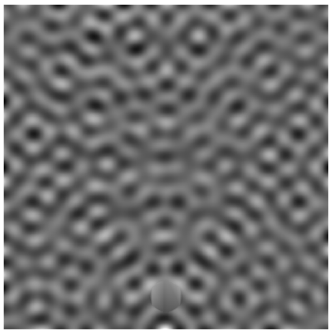
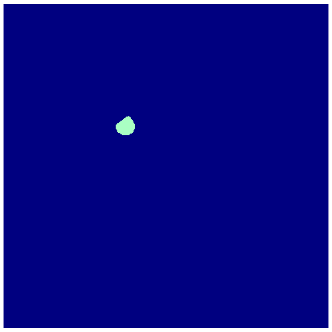
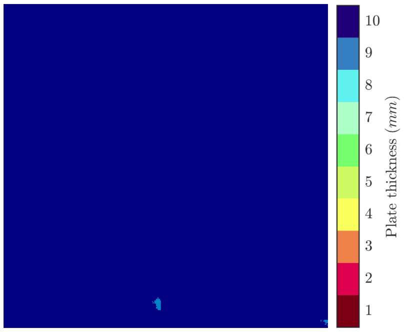
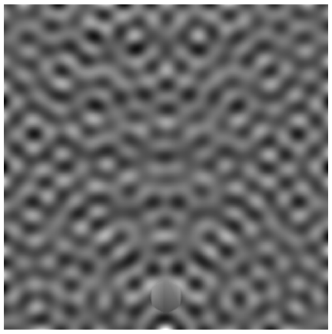
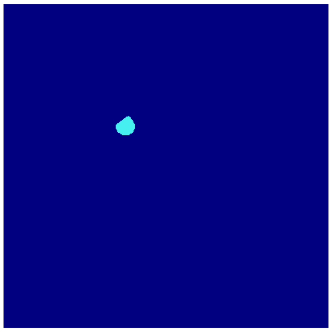
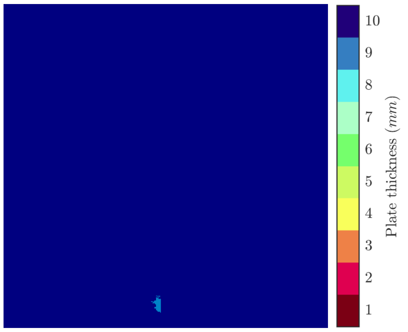
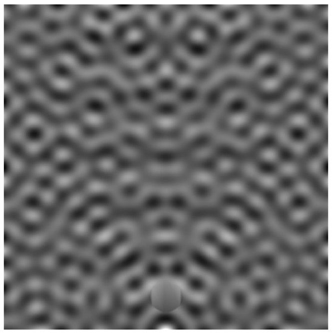
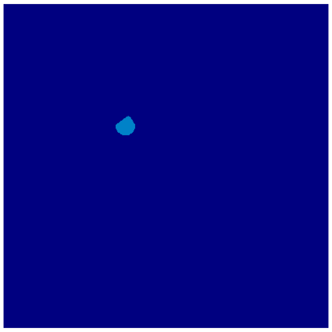
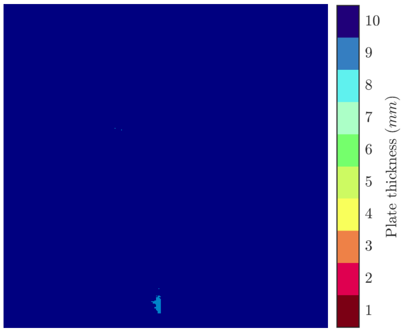
Appendices

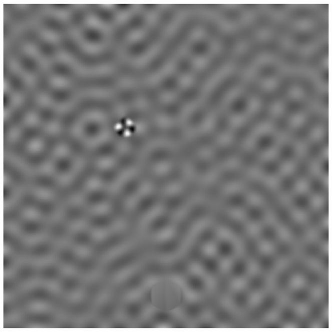
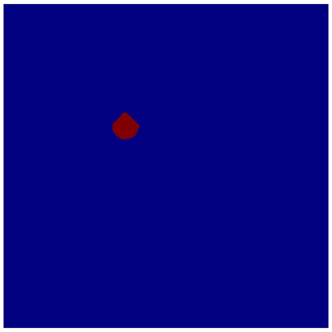
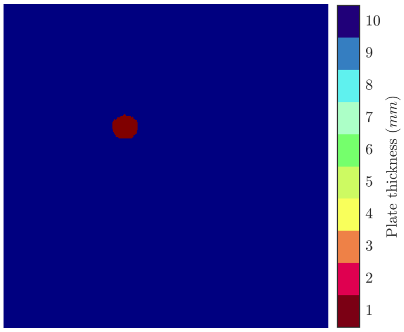
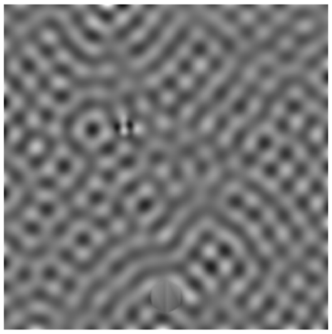
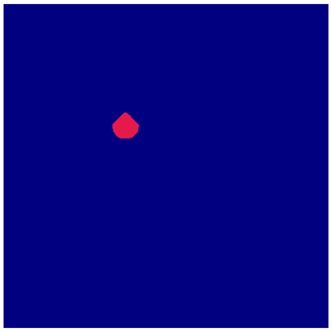
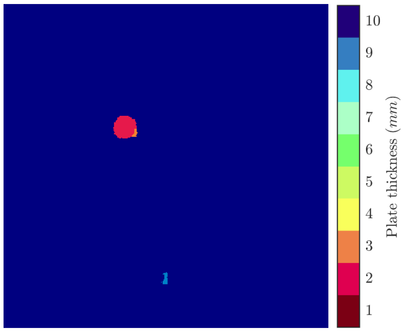
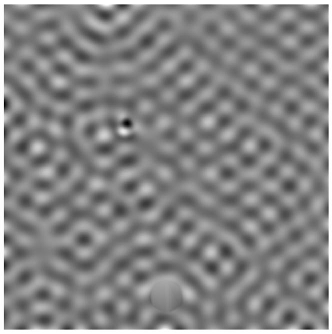
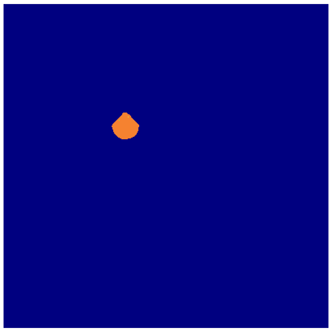
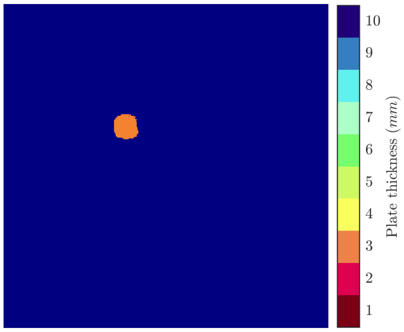
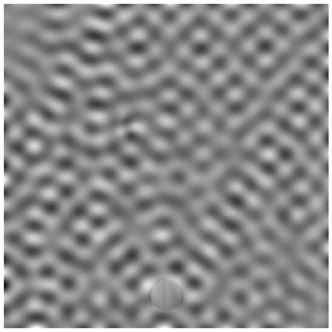
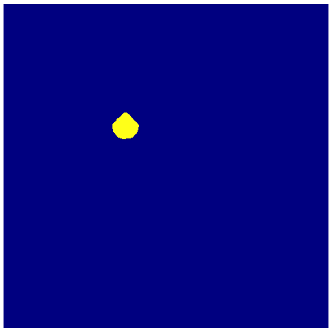
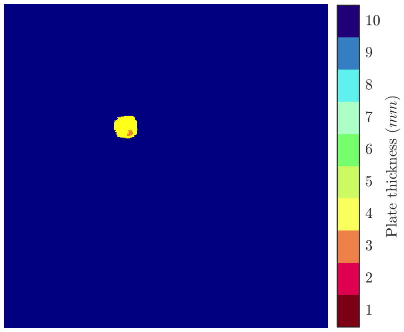
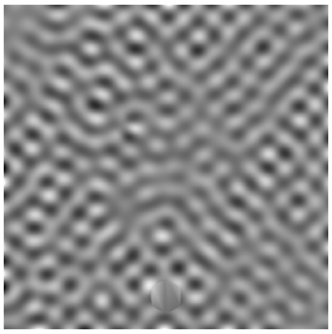
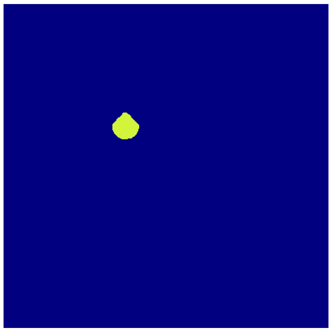
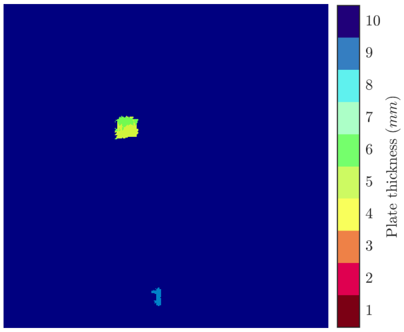
Appendix A. Full table of CNN prediction results on the simulation test set

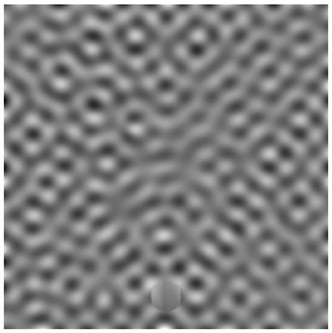
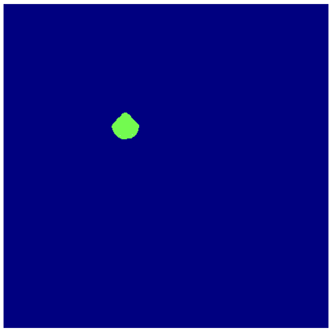
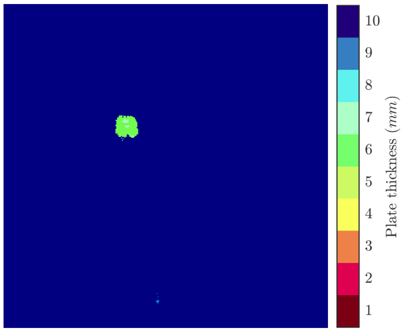
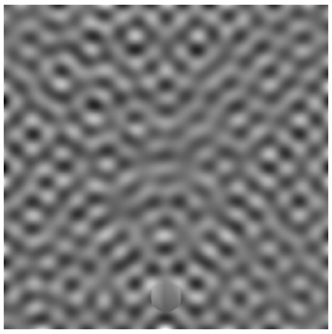
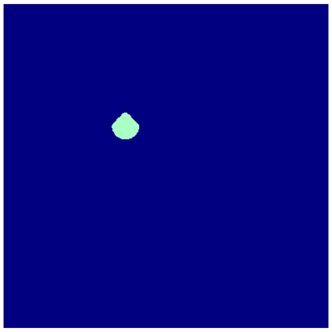
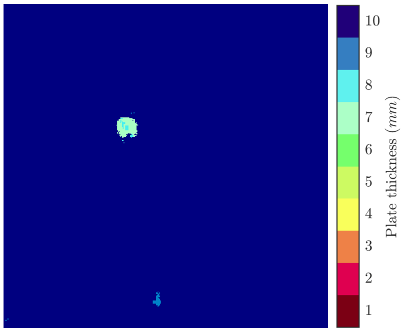
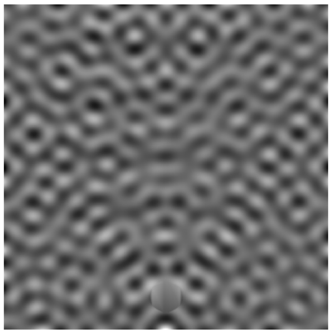
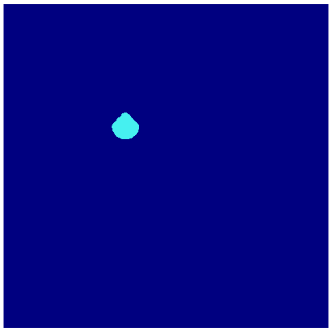
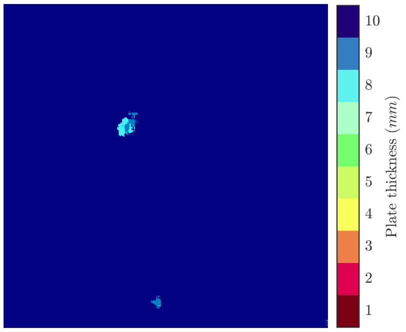
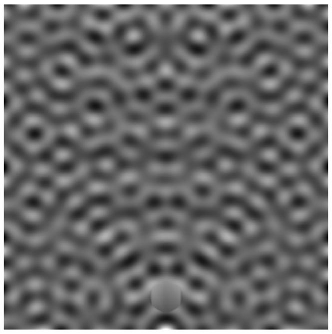
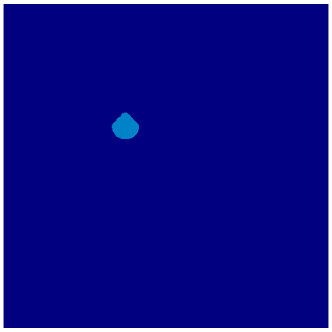
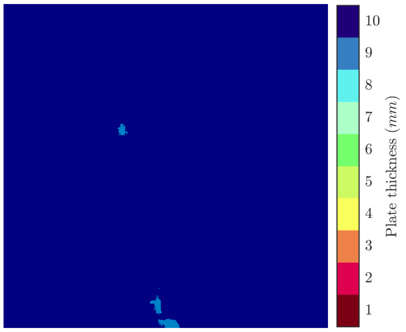
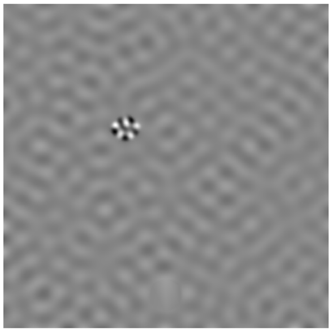
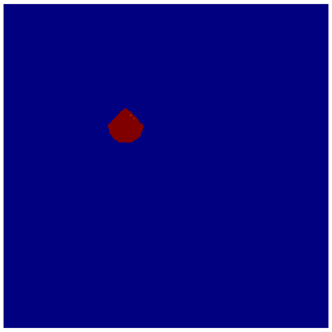
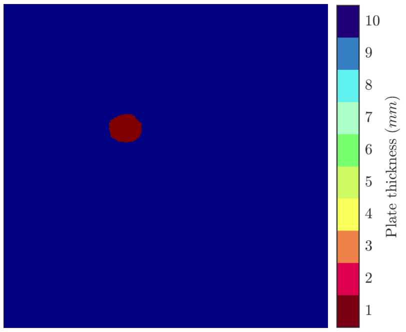
t	L	IOU	Wavefield	Ground Truth	Prediction
1	10	11%			
2	10	0%			
3	10	0%			

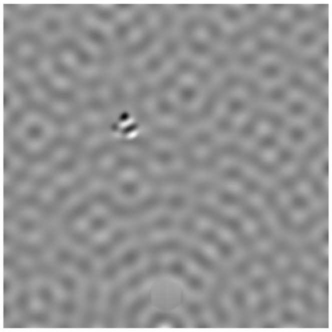
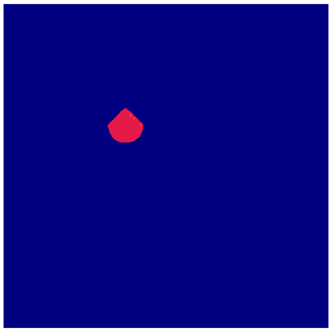
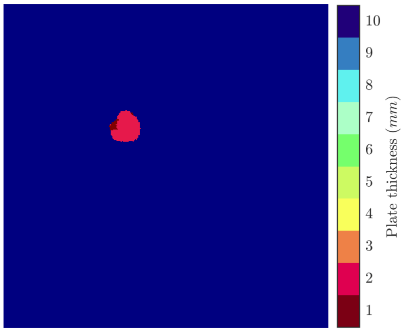
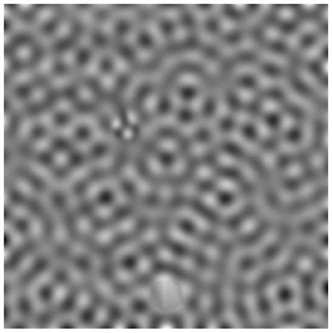
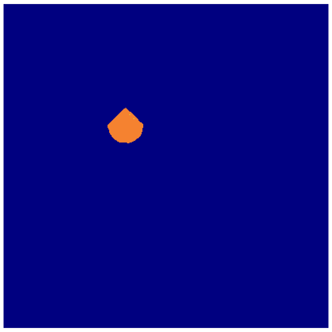
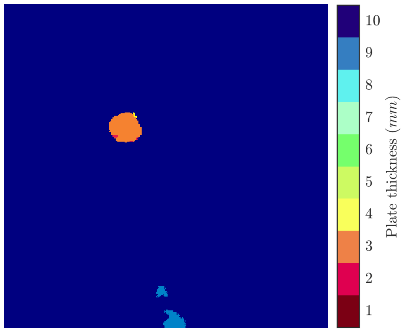
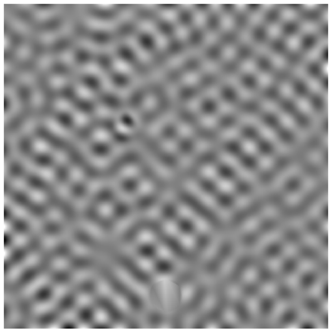
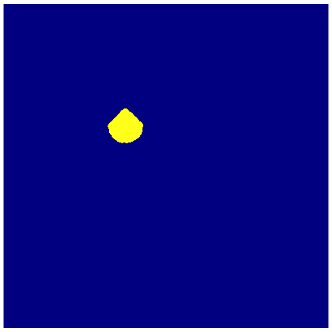
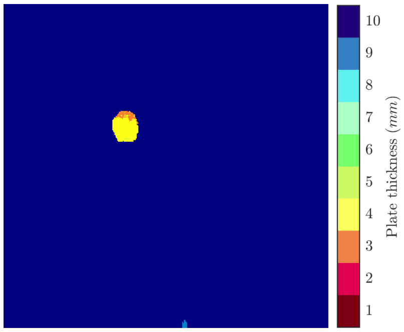
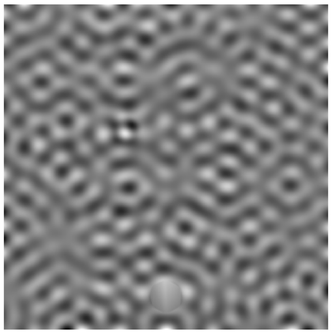
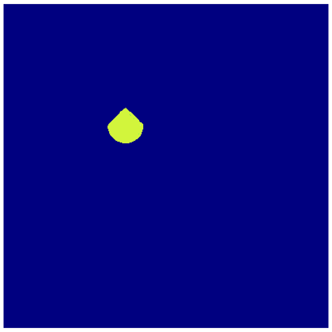
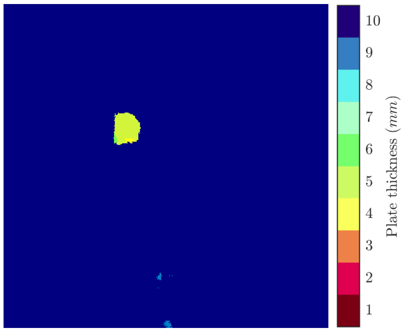
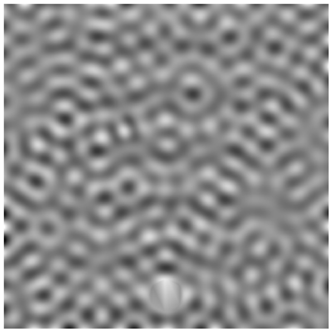
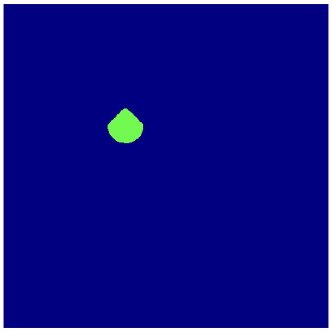
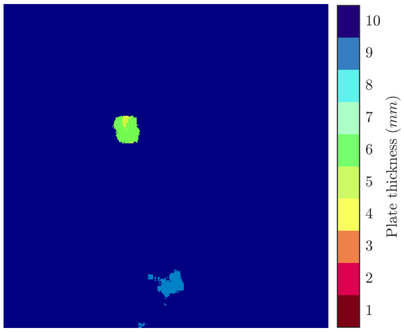
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
4	10	0%			
5	10	0%			
6	10	0%			
7	10	0%			
8	10	0%			

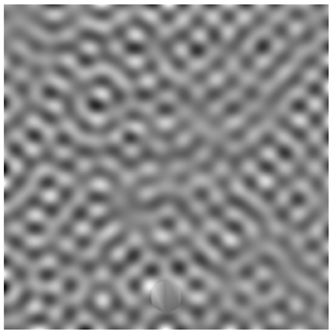
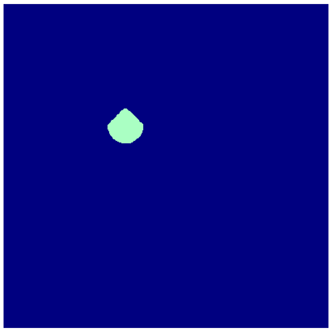
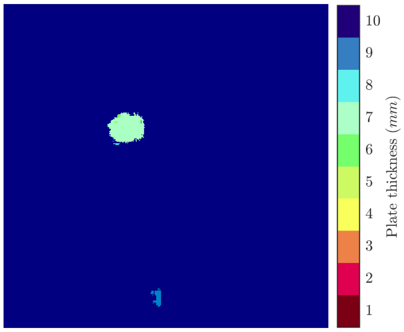
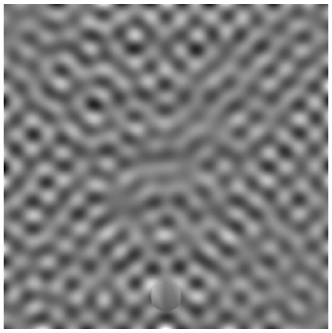
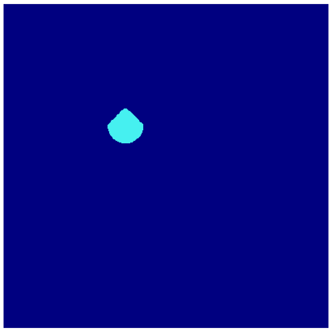
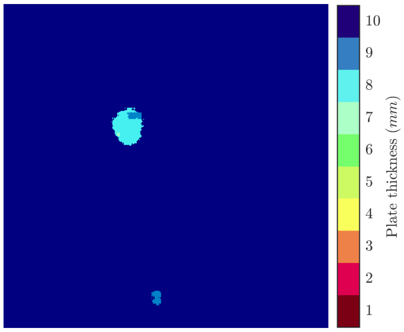
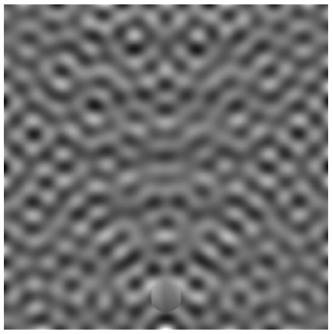
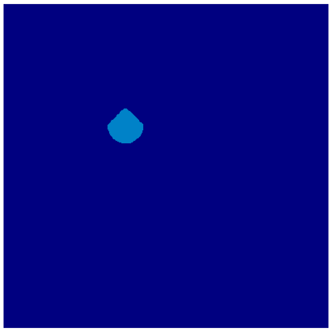
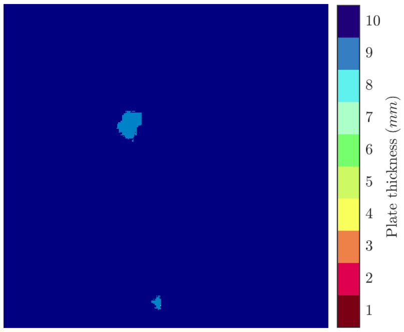
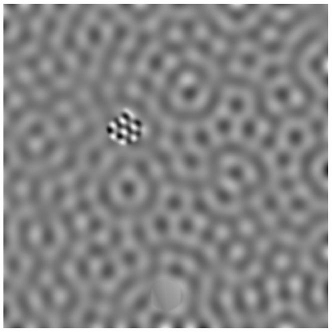
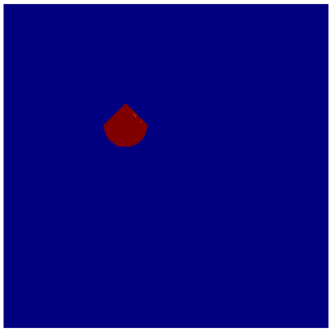
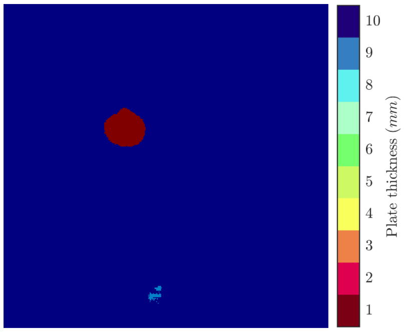
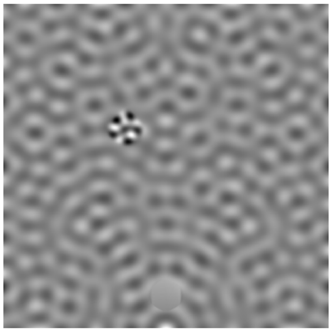
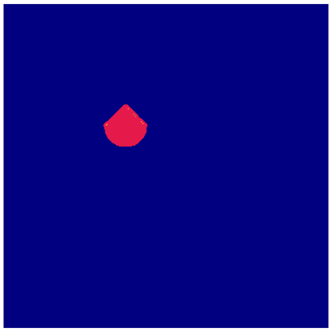
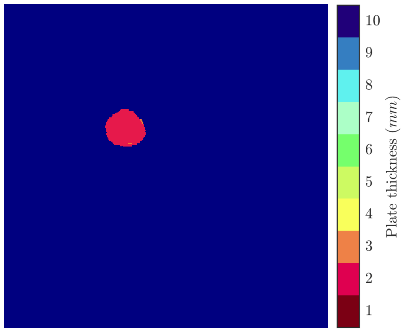
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
9	10	0%			
1	20	79%			
2	20	64%			
3	20	37%			
4	20	7%			

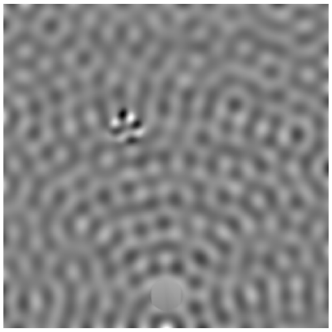
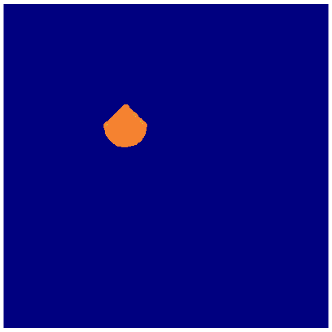
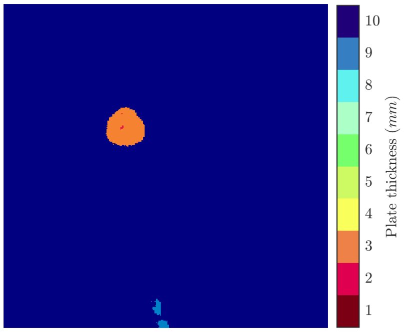
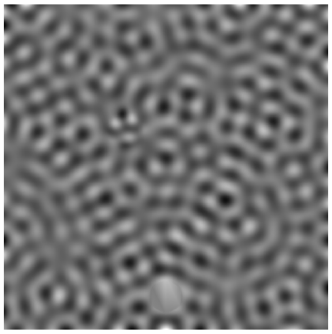
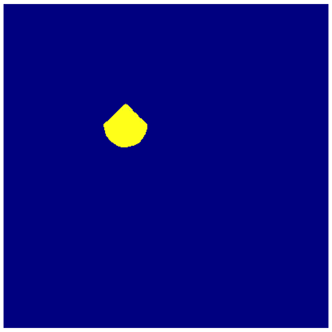
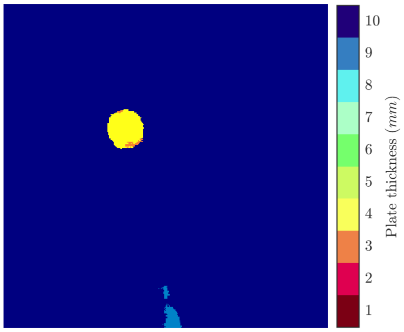
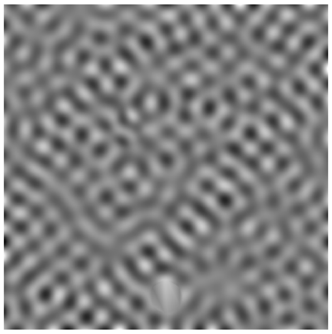
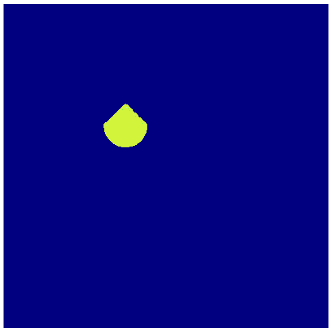
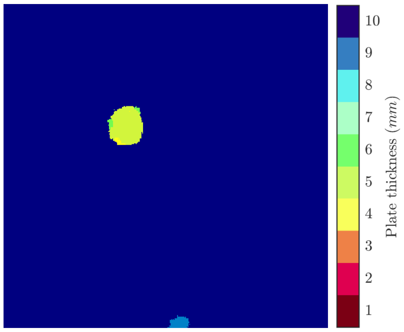
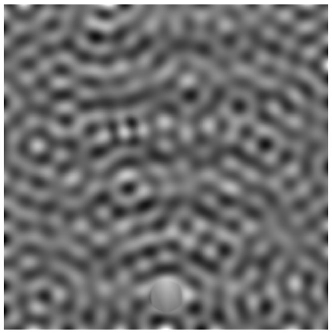
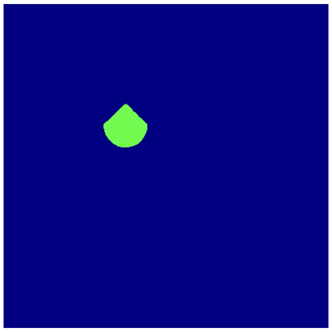
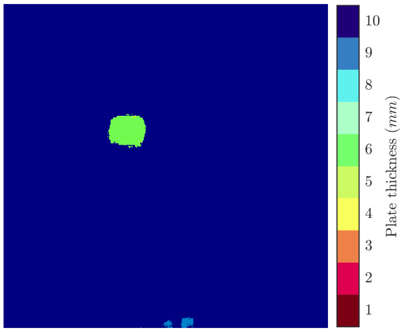
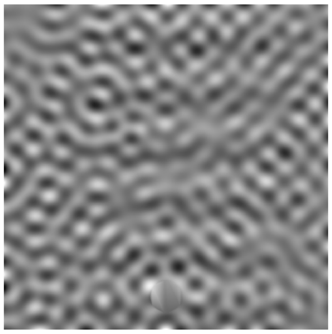
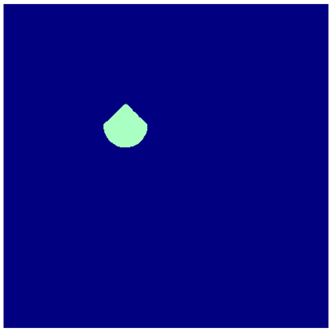
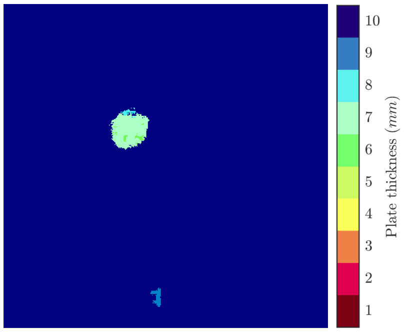
t	L	IOU	Wavefield	Ground Truth	Prediction
5	20	0%			
6	20	0%			
7	20	0%			
8	20	0%			
9	20	0%			

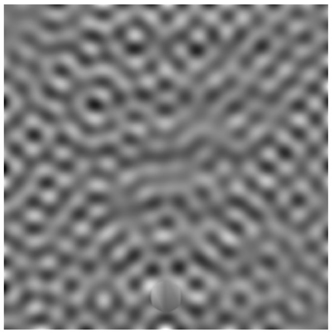
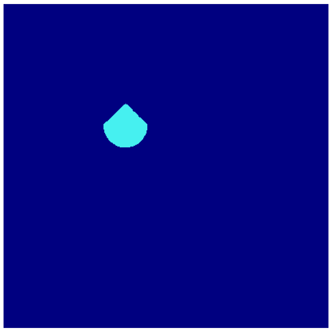
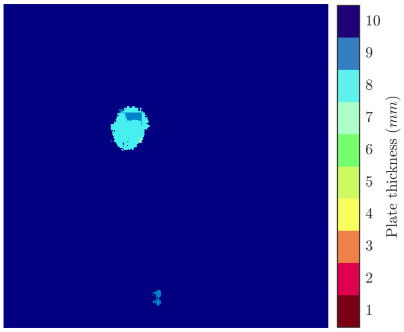
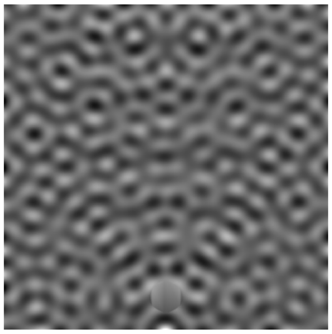
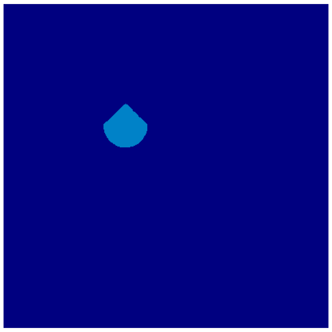
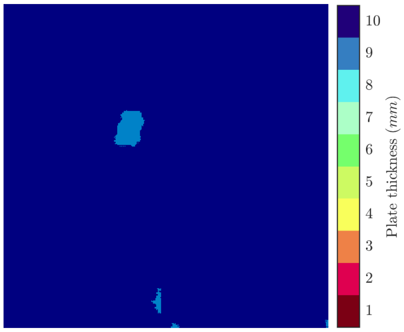
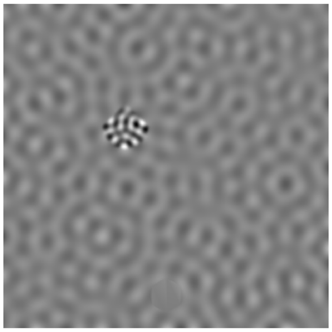
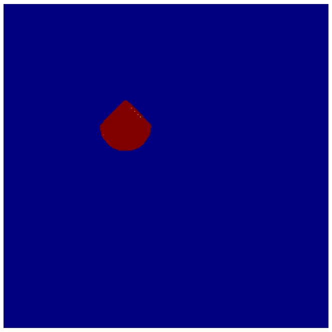
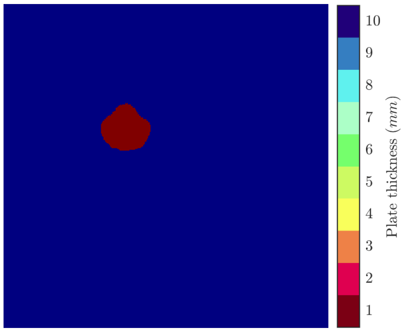
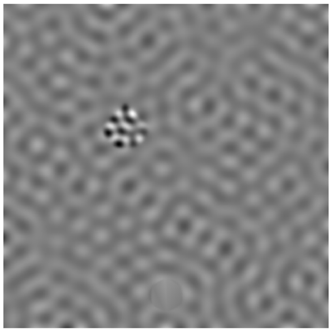
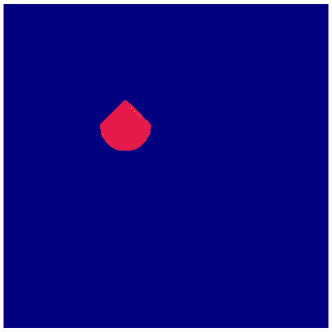
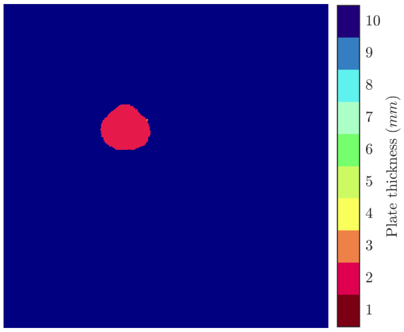
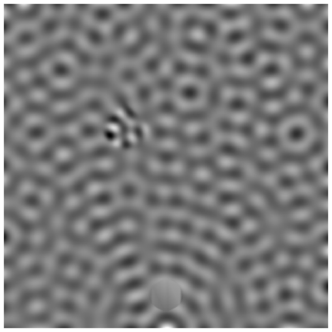
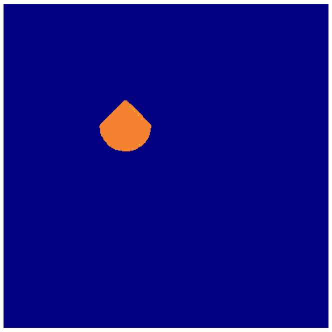
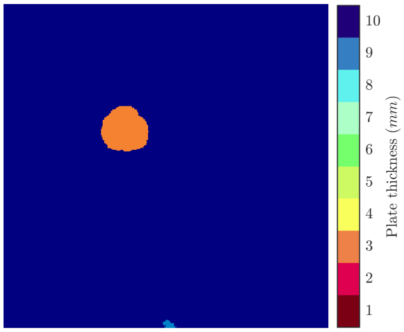
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
1	30	89%			
2	30	83%			
3	30	84%			
4	30	78%			
5	30	48%			

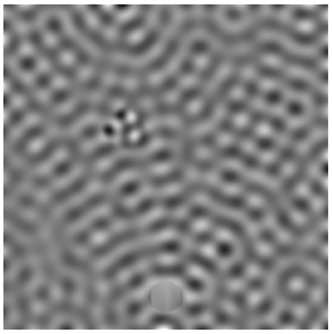
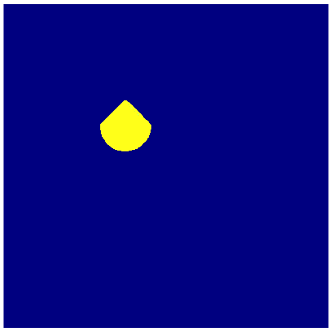
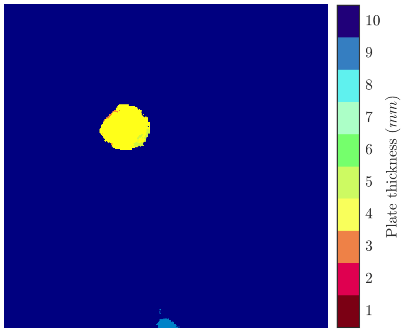
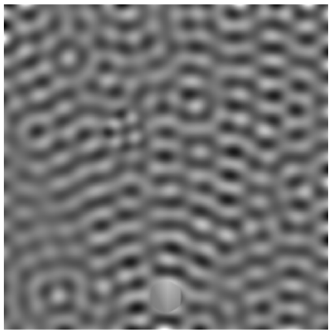
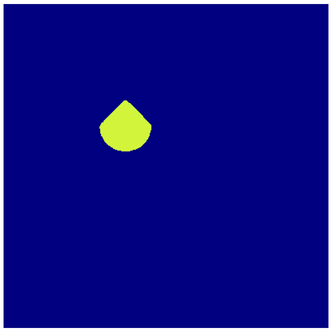
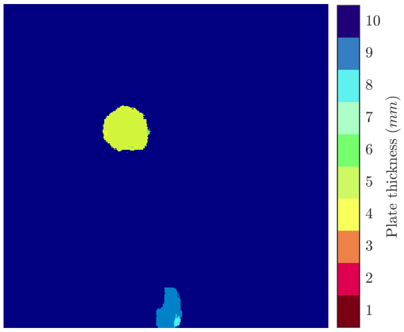
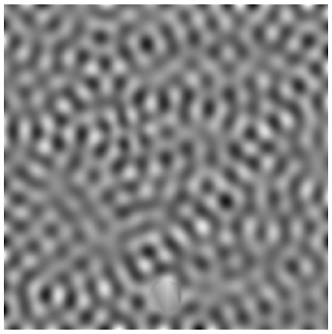
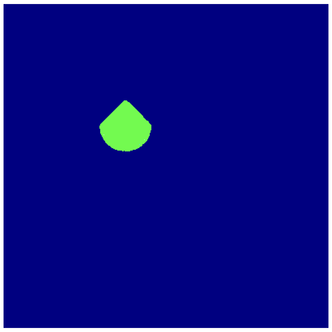
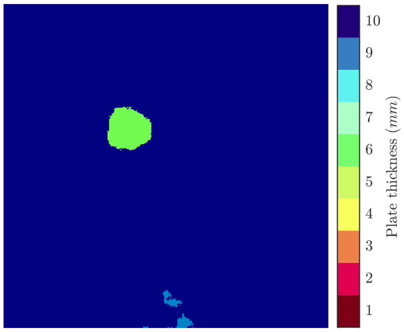
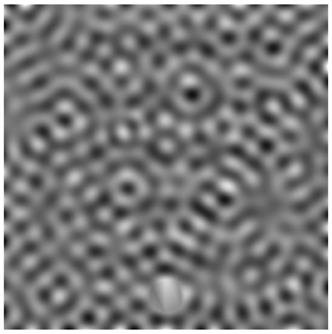
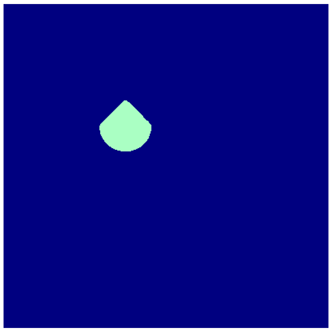
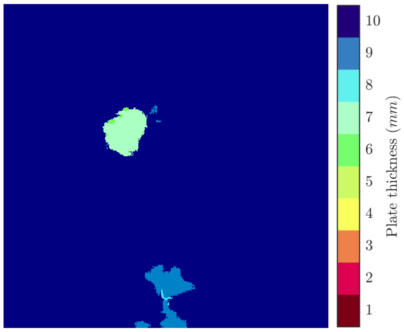
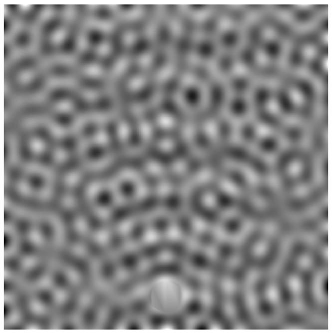
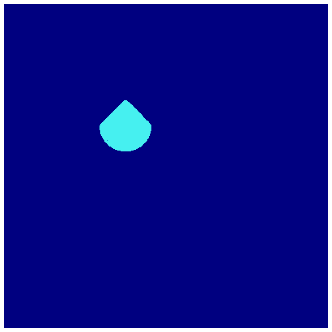
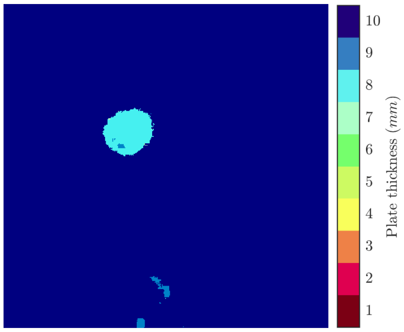
t	L	IOU	Wavefield	Ground Truth	Prediction
6	30	69%			
7	30	56%			
8	30	29%			
9	30	9%			
1	40	87%			

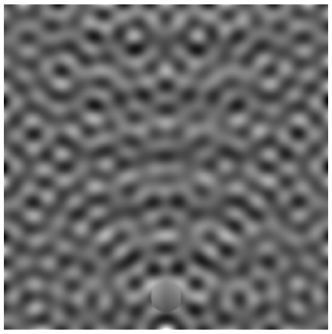
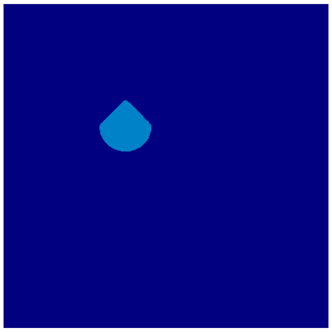
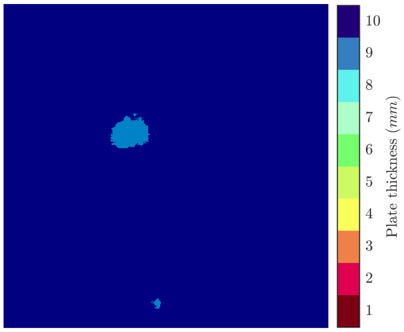
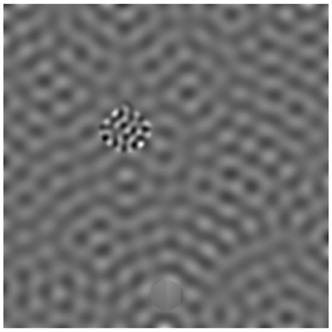
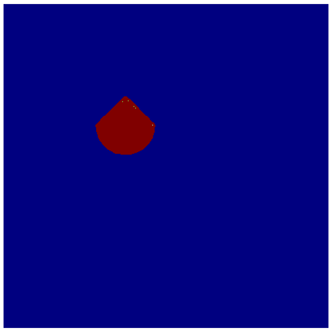
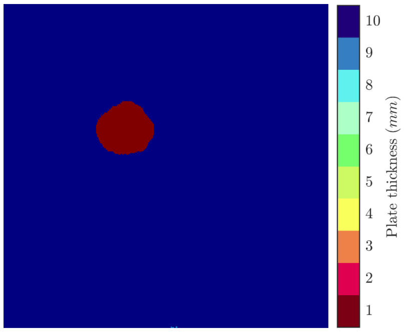
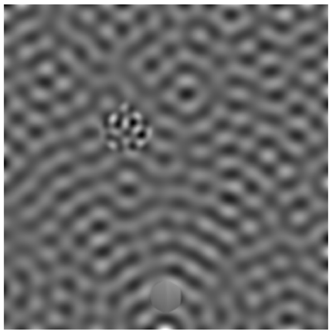
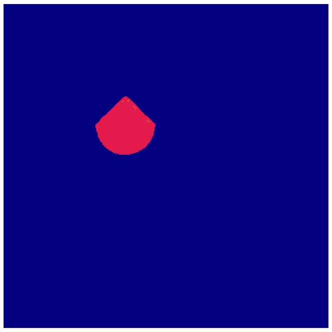
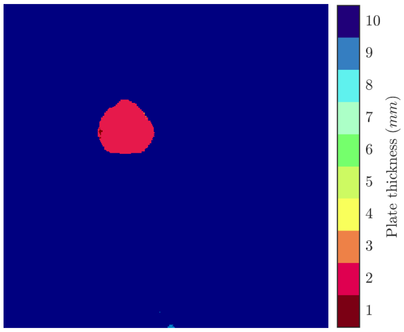
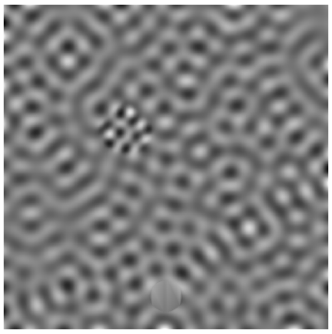
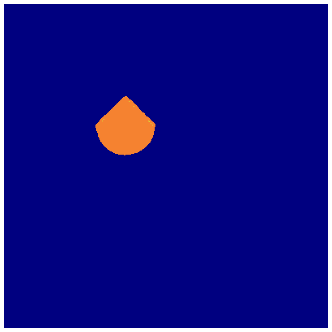
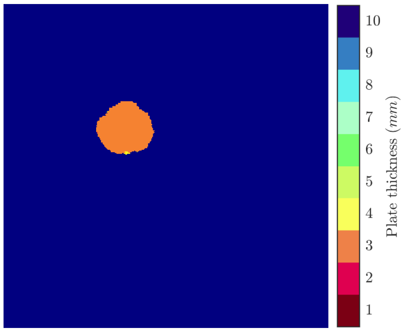
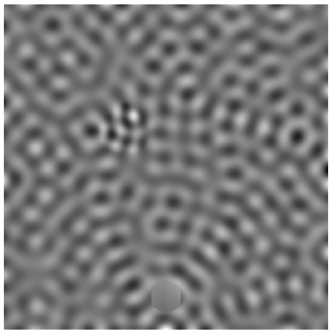
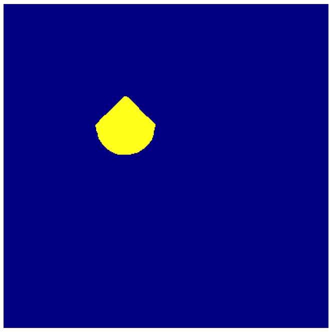
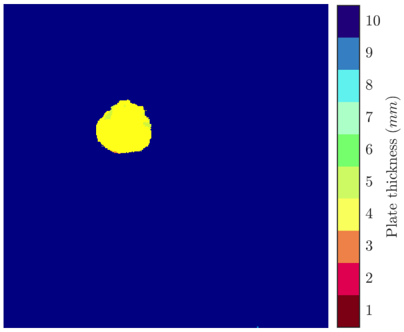
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
2	40	79%			
3	40	85%			
4	40	65%			
5	40	71%			
6	40	59%			

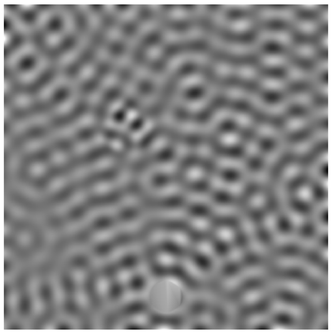
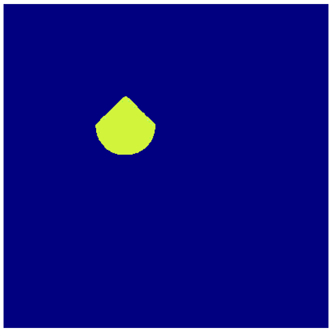
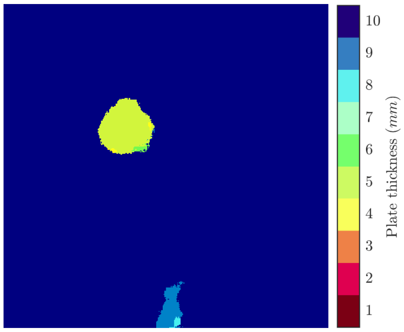
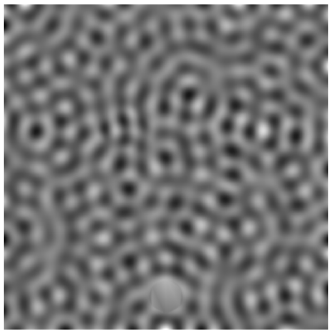
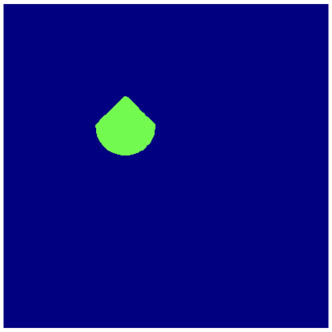
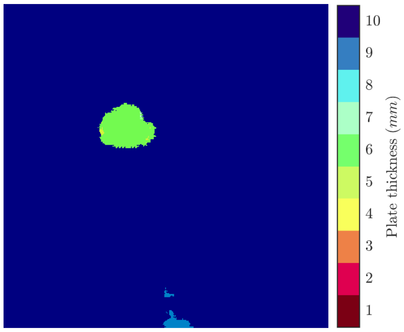
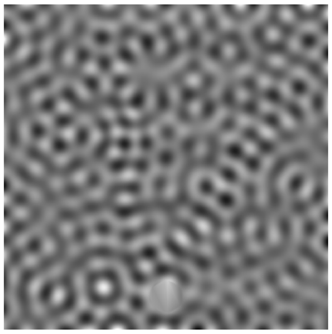
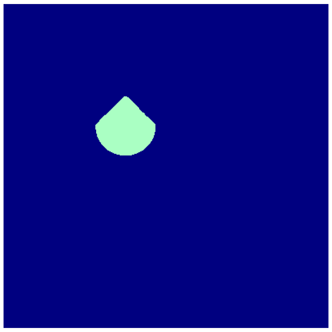
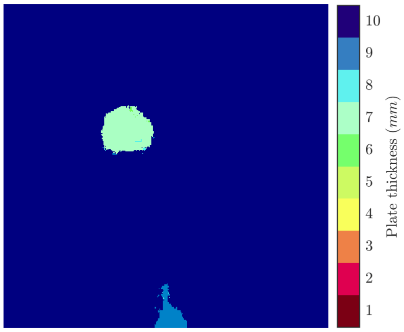
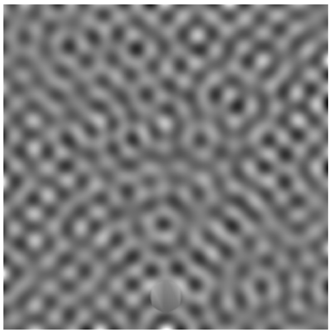
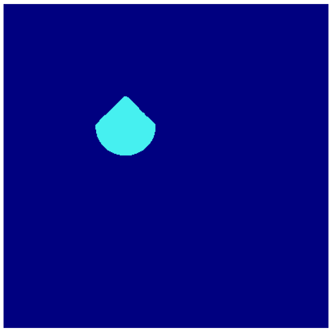
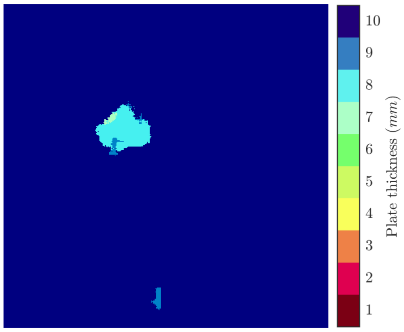
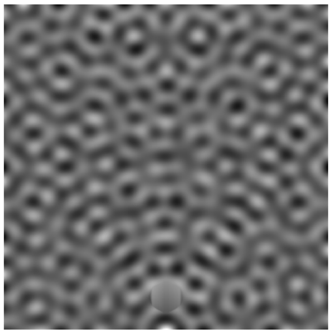
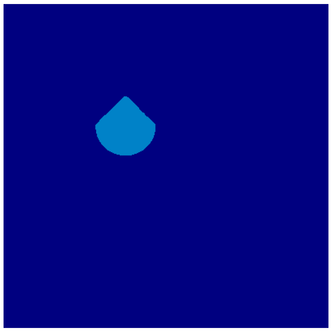
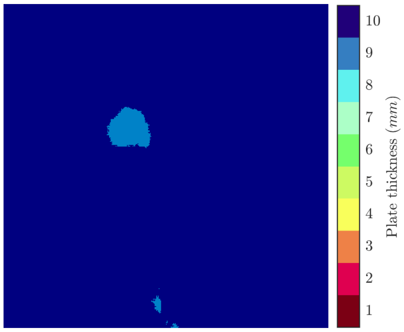
t	L	IOU	Wavefield	Ground Truth	Prediction
7	40	81%			
8	40	72%			
9	40	43%			
1	50	90%			
2	50	88%			

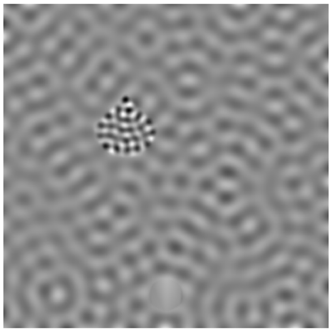
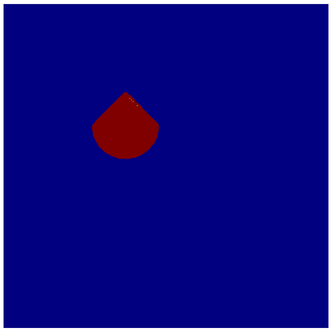
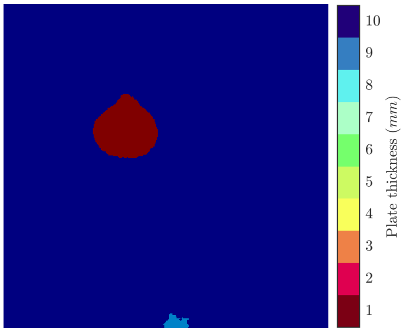
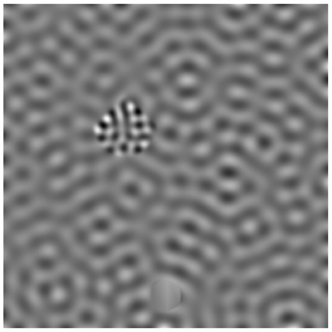
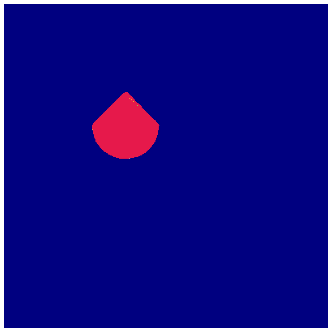
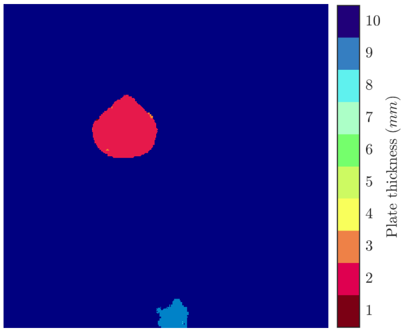
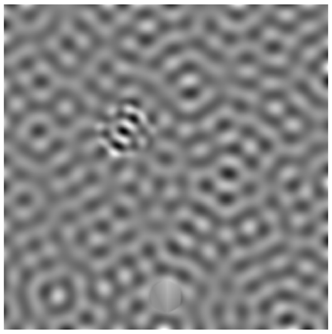
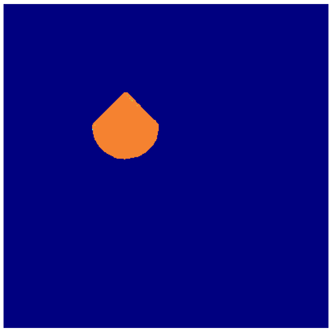
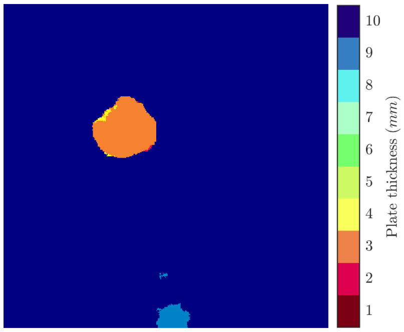
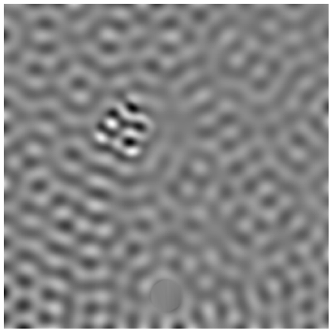
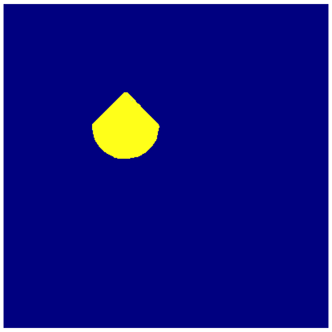
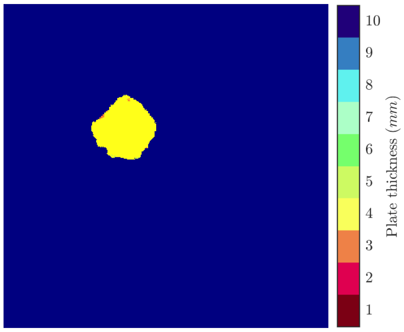
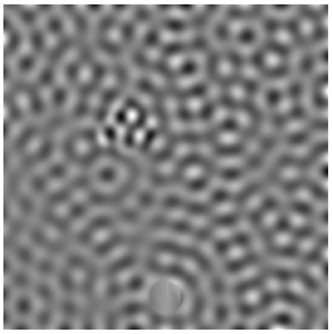
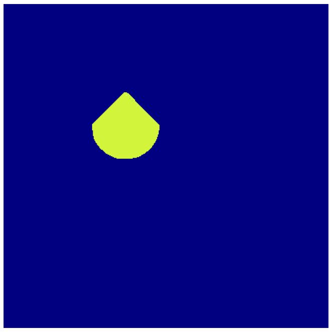
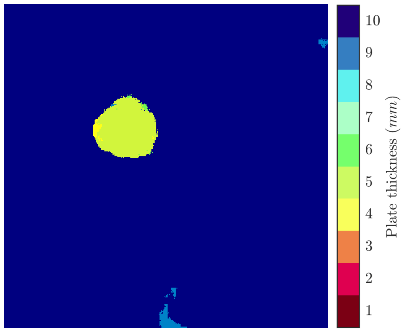
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
3	50	88%			
4	50	80%			
5	50	73%			
6	50	72%			
7	50	66%			

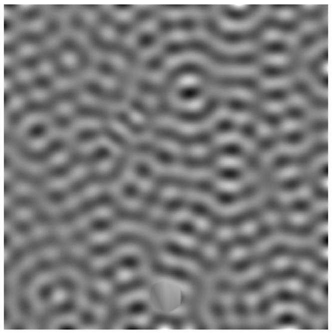
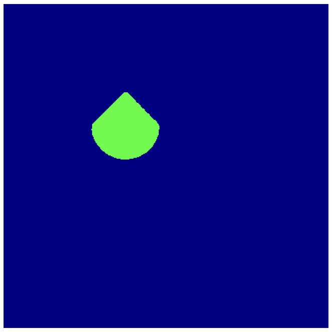
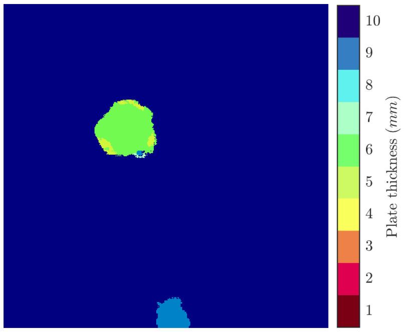
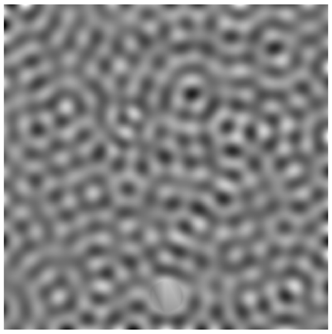
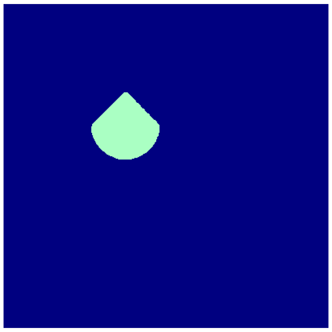
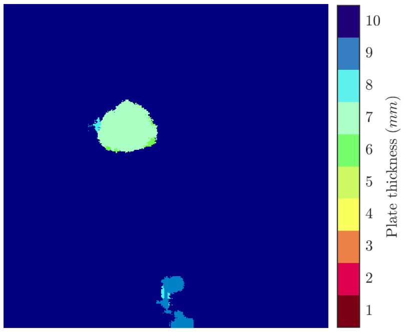
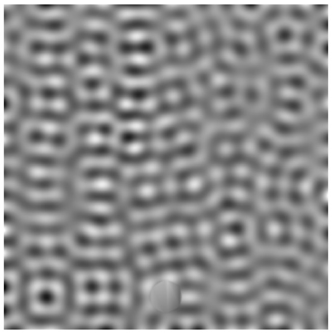
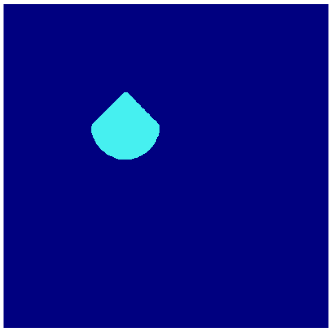
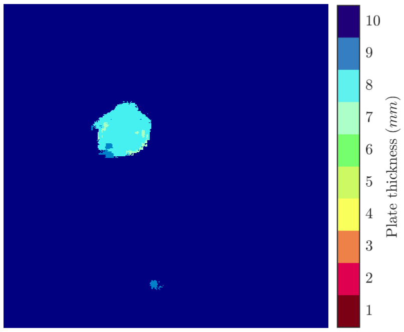
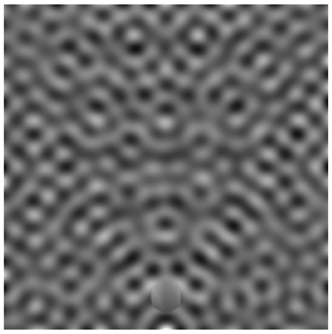
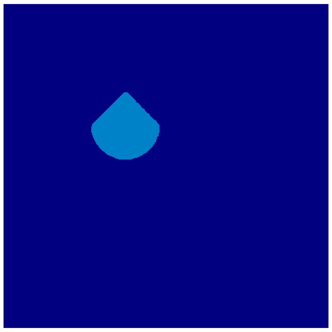
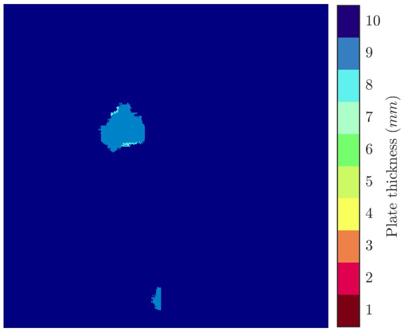
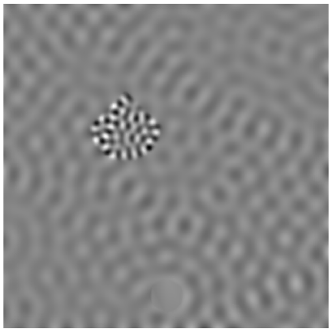
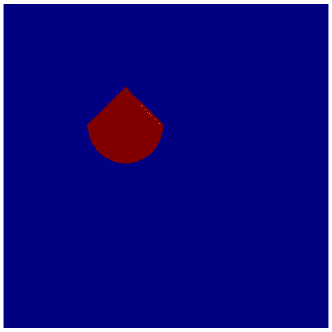
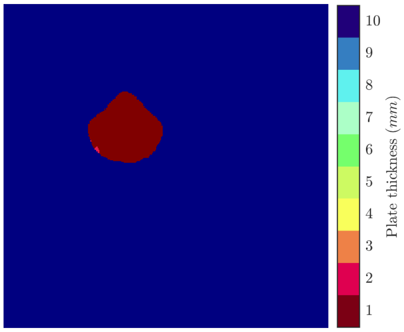
t	L	IOU	Wavefield	Ground Truth	Prediction
8	50	63%			
9	50	51%			
1	60	94%			
2	60	93%			
3	60	89%			

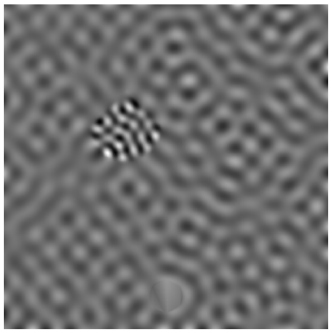
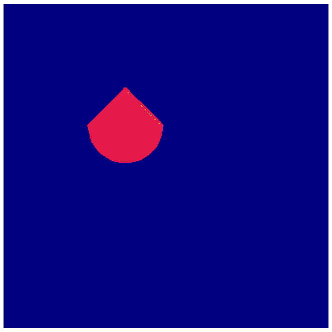
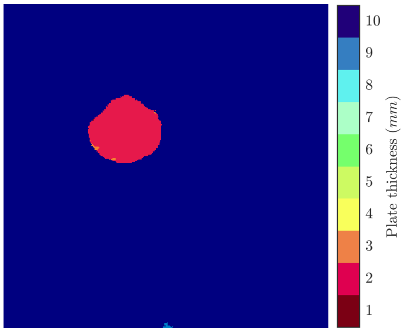
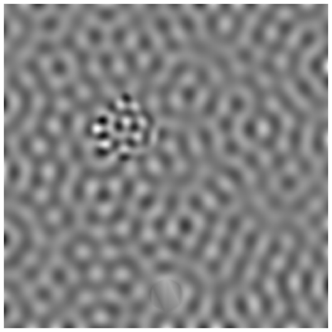
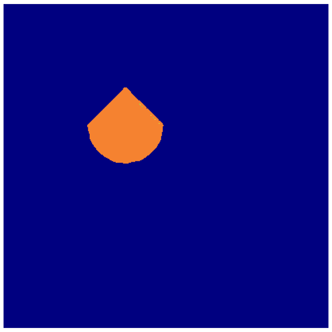
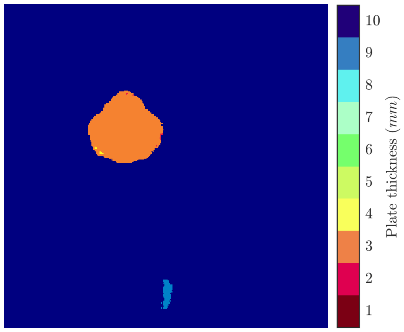
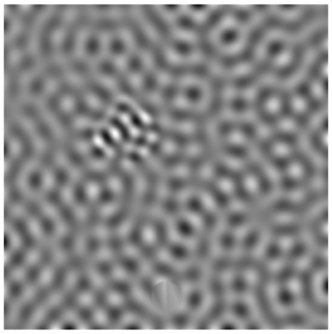
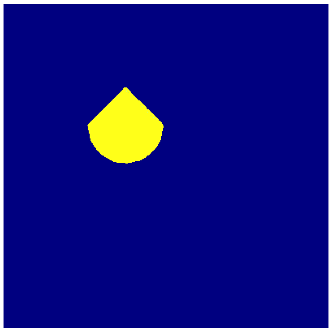
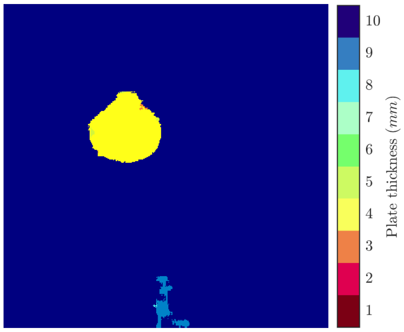
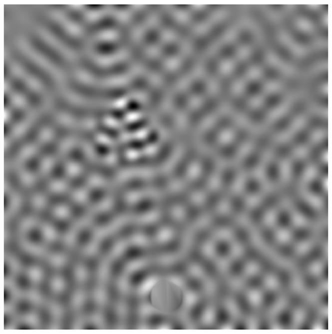
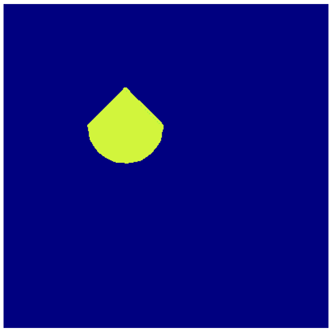
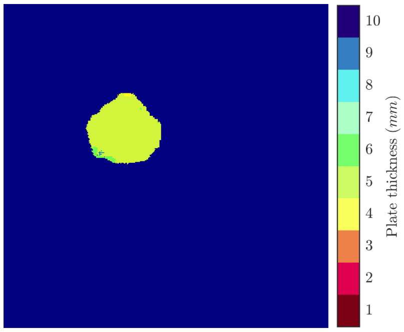
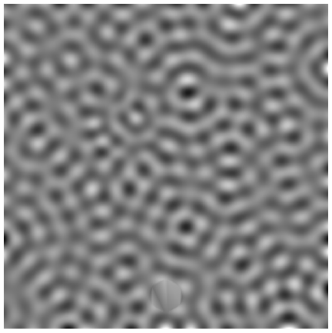
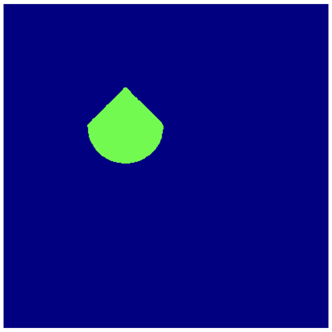
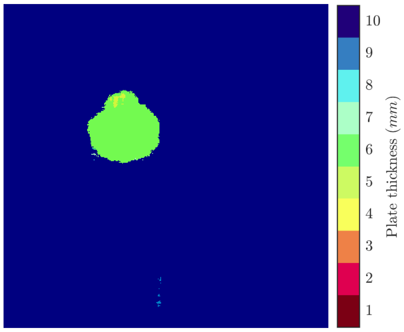
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
4	60	85%			
5	60	85%			
6	60	80%			
7	60	70%			
8	60	75%			

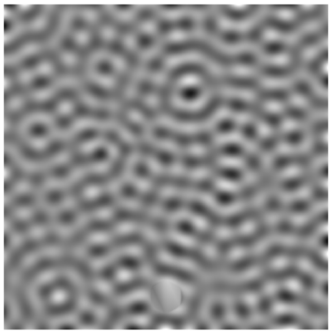
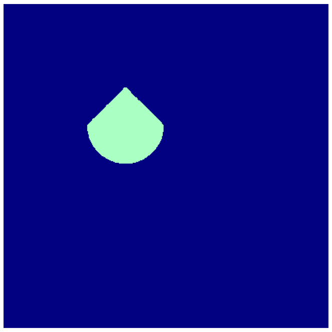
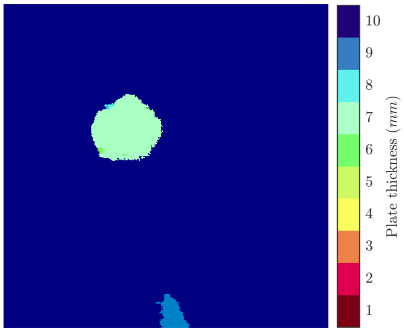
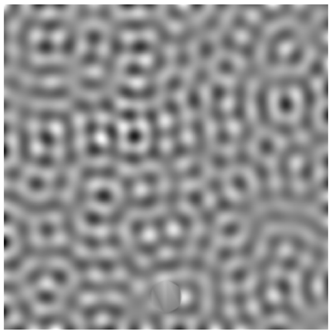
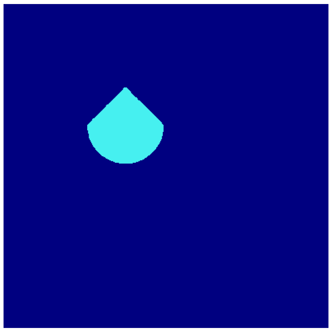
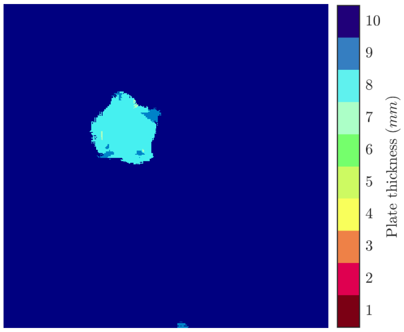
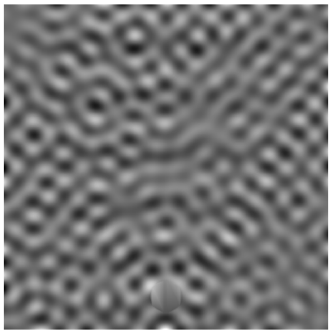
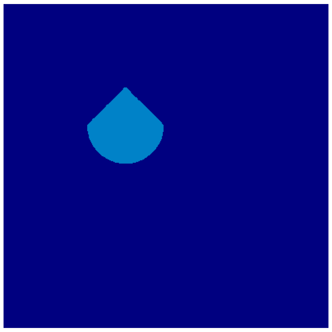
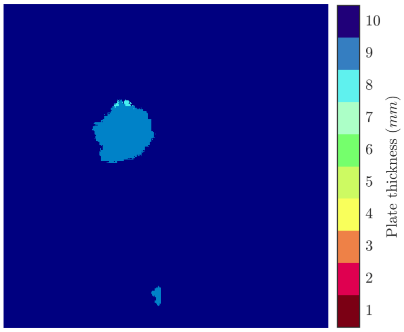
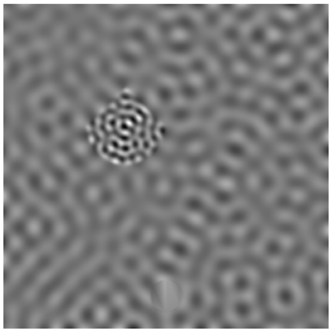
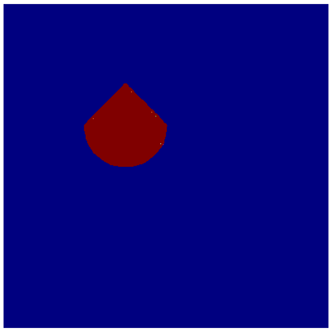
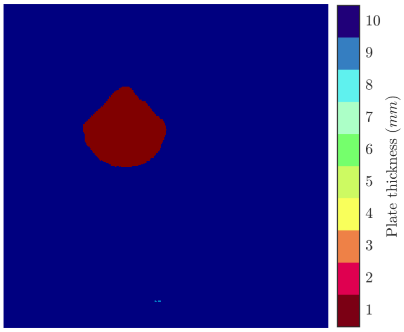
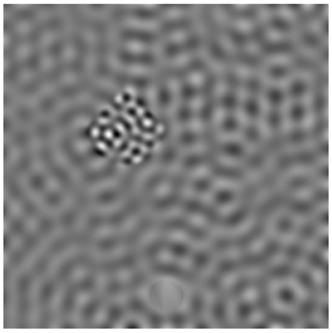
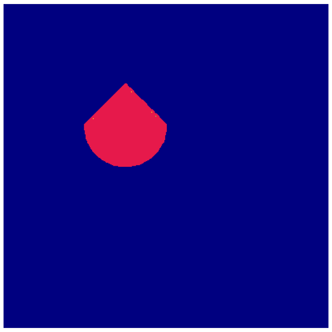
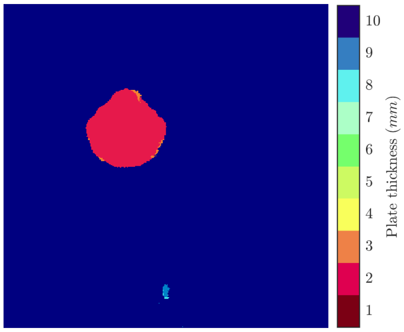
t	L	IOU	Wavefield	Ground Truth	Prediction
9	60	50%			
1	70	93%			
2	70	91%			
3	70	92%			
4	70	87%			

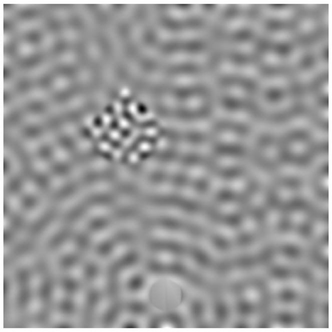
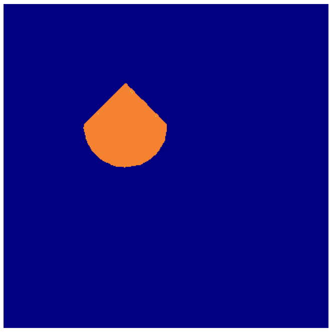
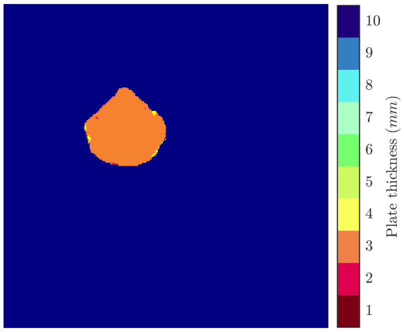
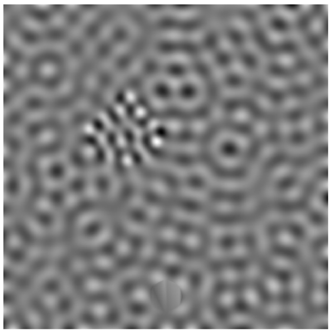
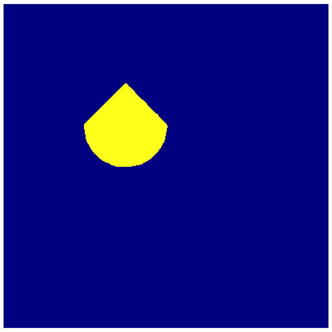
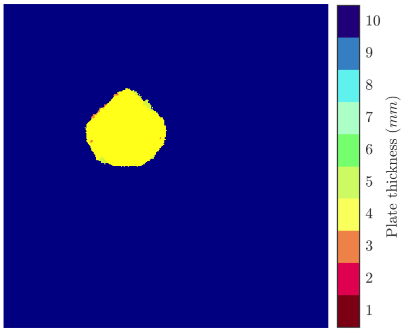
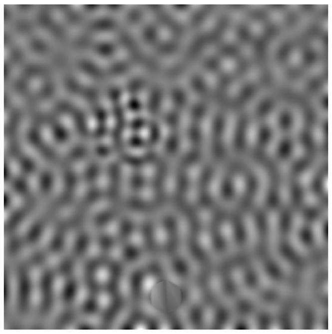
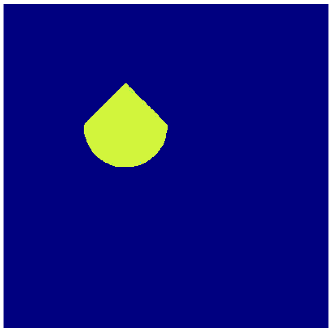
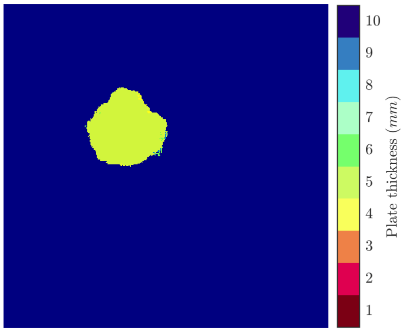
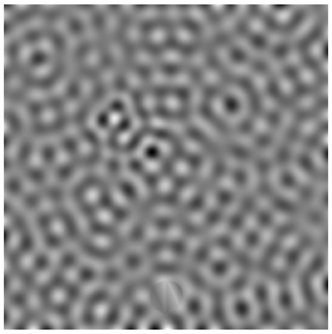
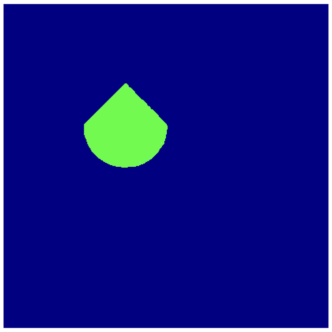
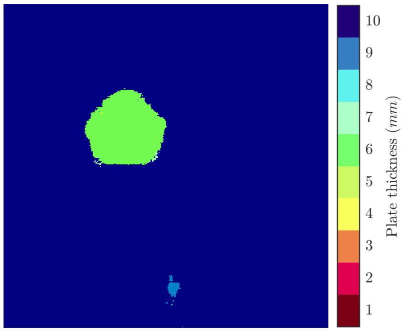
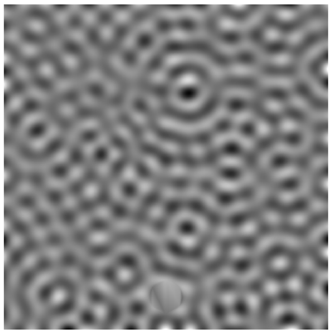
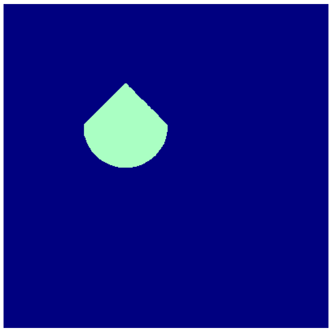
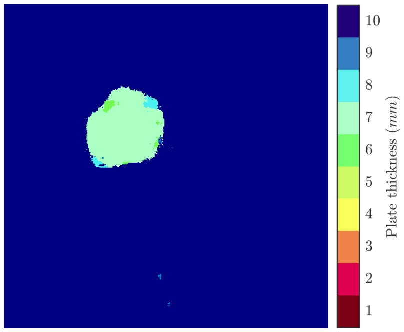
t	L	IOU	Wavefield	Ground Truth	Prediction
5	70	86%			
6	70	75%			
7	70	78%			
8	70	63%			
9	70	52%			

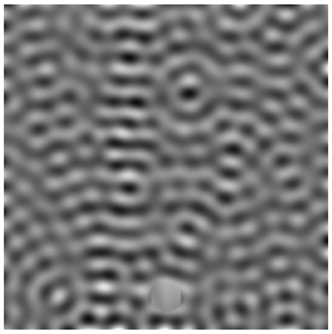
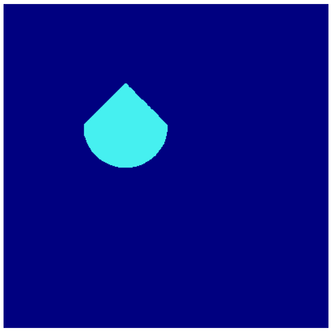
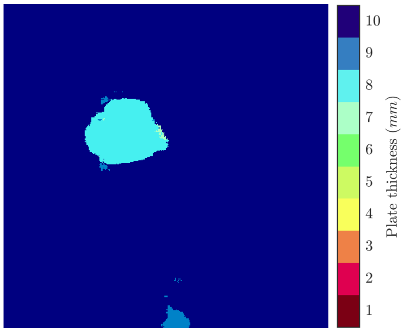
t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
1	80	95%			
2	80	94%			
3	80	90%			
4	80	90%			
5	80	88%			

t	L	IOU	Wavefield	Ground Truth	Prediction
6	80	73%			
7	80	71%			
8	80	67%			
9	80	43%			
1	90	96%			

t	L	IOU	Wavefield	Ground Truth	Prediction
2	90	93%			
3	90	94%			
4	90	93%			
5	90	91%			
6	90	91%			

t	L	IOU	Wavefield	Ground Truth	Prediction
7	90	86%			
8	90	78%			
9	90	64%			
1	100	96%			
2	100	95%			

t	*L*	*IOU*	Wavefield	Ground Truth	Prediction
3	100	94%			
4	100	93%			
5	100	91%			
6	100	91%			
7	100	85%			

t	L	IOU	Wavefield	Ground Truth	Prediction
8	100	81%			
9	100	84%