Machine Learning for Improved Current Density Reconstruction from 2D Vector Magnetic Images

Niko R. Reed,[1,2,3,*] Danyal Bhutto,[4,5,*] Matthew J. Turner,[1,2,6] Declan M. Daly,[1,2] Sean M. Oliver,[7] Jiashen Tang,[1,2] Kevin S. Olsson,[2,6,8] Nicholas Langellier,[1,9,†] Mark J.H. Ku,[10,11,‡] Matthew S. Rosen,[5,9,12] and Ronald L. Walsworth[1,2,6,§]

[1] *Department of Physics, University of Maryland, College Park, MD, USA*
[2] *Quantum Technology Center, University of Maryland, College Park, MD, USA*
[3] *Joint Quantum Institute, University of Maryland, College Park, MD, USA*
[4] *Department of Biomedical Engineering, Boston University, Boston, MA, USA*
[5] *A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, MA, USA*
[6] *Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA*
[7] *The MITRE Corporation, McLean, VA 22102, United States of America*
[8] *Intelligence Community Postdoctoral Research Fellowship Program, University of Maryland, College Park, Maryland 20742, USA*
[9] *Department of Physics, Harvard University, Cambridge, MA, USA*
[10] *Department of Materials Science and Engineering, University of Delaware, Newark, DE 19716, USA*
[11] *Department of Physics and Astronomy, University of Delaware, Newark, DE 19716, USA*
[12] *Harvard Medical School, 25 Shattuck St., Boston, MA, USA*

(Dated: January 11th, 2025)

The reconstruction of electrical current densities from magnetic field measurements is an important technique with applications in materials science, circuit design, quality control, plasma physics, and biology. Analytic reconstruction methods exist for planar currents, but break down in the presence of high spatial frequency noise or large standoff distance, restricting the types of systems that can be studied. Here, we demonstrate the use of a deep convolutional neural network for current density reconstruction from two-dimensional (2D) images of vector magnetic fields acquired by a quantum diamond microscope (QDM). Trained network performance significantly exceeds analytic reconstruction for data with high noise or large standoff distances. This machine learning technique can perform quality inversions on lower SNR data, significantly reducing the data collection time and permitting reconstructions of weaker and three-dimensional current sources.

I. INTRODUCTION

The study of electric current distributions through imaging of Oersted magnetic fields has supported applications in many areas of physics, engineering, and medicine, in part due to its non-invasive nature [1]. Such applications include assessment of integrated circuits [1–7], microelectronics [8–10], batteries [11–14], solar panels [15, 16], superconducting tape and wires [17, 18], mechanical joints [19], fault analysis [20], superconducting materials and qubits [21–23], exotic materials [24–27], new semiconductor technology [28, 29], and eddy current imaging [30]. In medicine, imaging of magnetic fields generated by currents in biological tissues allows detailed and unique study of the heart, brain, and skeletal muscles [31–37].

To enable these studies, current density distributions confined to thin planes are traditionally calculated utilizing Fourier space relationships between the source current density and the measured magnetic field. Despite its widespread usage, this "Fourier Method" for the magnetic inverse problem (i.e. the current density reconstruction problem) cannot effectively address high-noise measurements [38]; it also performs poorly when the standoff distance between the measured magnetic field and the source current density is greater than the feature size of the current distribution. Due to these limitations, the Fourier Method sometimes results in poor analytic reconstructions [7, 38]. These drawbacks most heavily impact its utility for systems with small features or restricted standoff distance, such as nondestructive quality control and failure analysis of intact semiconductor chips, imaging of wires through walls or device boundaries, and noninvasive measurements of biological systems, which commonly exhibit weak magnetic field signals and small, irregular features [31].

Due to the poor reconstruction performance of the Fourier Method in the low signal-to-noise ratio (SNR) regime and/or at large standoff distance, many repeated magnetic field measurements are typically required to achieve sufficient SNR, significantly slowing the measurement process and leading to systematic errors related to

* Indicates co-first authors
† Present address: VideaHealth Inc., Boston, MA
‡ Present address: Northrop Grumman Mission Systems, Linthicum, Maryland 21090, USA.
§ walsworth@umd.edu

drift in the measured system or instrumentation [39]. An alternative current reconstruction technique requiring fewer measurements could enable higher fidelity results, the study of dynamic processes in samples that would otherwise be immeasurable, and greatly increase the speed of device quality control applications.

Neural networks excel at tackling a variety of image-based inverse problems [40–43]. While early attempts at applying neural networks to the magnetic inverse problem operated in highly restricted parameter spaces [44, 45], advances in computing power and machine learning techniques show promise for identification of current density features of unrestricted sizes, shapes, and strengths in noisy images [46].

Recently, a physically-informed network, trained to find the current density distribution that best reproduces the input magnetic field, has demonstrated success in low noise regimes, and can even can outperform the Fourier Method [47, 48]. While this approach has the benefit of not requiring independent training data, it is not suitable for high noise inputs, as the network will seek to reproduce all input features including noise. A network trained with data pairs of noisy magnetic field inputs and associated current density distribution targets could allow the network to learn to differentiate signal from noise in the magnetic input signal. In particular, convolutional neural networks, including UNets [49] have shown impressive results in medical imaging reconstruction problems with regard to de-noising and artifact reduction [50, 51]; and more generally in preserving structural features [40]. In addition to the demonstrated performance of UNet architectures, using a fully convolutional network reduces the computational demand of training with large images, and produces a network which can accept images of multiple resolutions. Both advantages stem from the fact that convolutional layers act by convoluting a filter kernal with the input data. Unlike linear layers, the size of the filter kernal is independent of the size of the input, allowing flexibilty in inference and enabling training with higher resolution data.

In this work, we present an adapted UNet architecture (MAGnetic Inverse Calculation UNet) designed for reconstructing current density images from vector magnetic field images with greater accuracy than the Fourier Method. The MAGIC-UNet network is tested on a variety of synthetic data types and experimentally collected Oersted magnetic field images using a Quantum Diamond Microscope (QDM). The QDM utilizes a dense, near-surface layer of nitrogen vacancy (NV) quantum defects in a diamond substrate to enable sensitive magnetic imaging, with micron-scale spatial resolution, via optical emission of the NV defects [39]. The MAGIC-UNet network is trained, in less than twelve hours, on hundreds of thousands of field-current image pairs generated analytically through Biot-Savart forward calculations of vector magnetic fields from randomly generated planar current density distributions (Figure 1a). We present current density reconstructions from both simulations and QDM experiments utilizing the MAGIC-UNet network, demonstrating greatly enhanced performance compared to the Fourier Method, and reliable reconstructions even in the presence of extreme noise (input SNR ~ 1).

II. NETWORK ARCHITECTURE

Characteristic to UNet architectures, MAGIC-UNet combines low and high resolution image filtering through a series of convolutional layers [49]. The purpose of the network is to take a vector magnetic field image in a plane above the sample (expressed in components B_x, B_y and B_z for each pixel in the image), and output the associated planar current density distribution (components J_x and J_y) for each point in the source plane. The network architecture is divided into a downsampling portion followed by a branched upsampling portion, with one branch for each component of current density.

The MAGIC-UNet network begins image processing with a series of downsampling steps that reduce the image size and allow the network to capture low-level features [40]. During this process, the resolution of a given input image is reduced by a factor of 2 for each step while the number of channels is increased by 2. After 3 downscaling iterations, a similar upsampling process occurs that increases the resolution, ultimately producing final outputs of the same resolution as the starting images. In addition, layers from the downsampling path are directly connected through skip connections to upsampled layers of the same resolution to bypass intermediate layers. These connections enable the preservation of structure in the data across multiple resolutions that would otherwise be lost due to downsampling. Skip connections also ameliorate the vanishing gradient problem by allowing gradients to flow to earlier layers before the gradient diminishes [52]. We omit the first skip connection as we find it does not impact MAGIC-UNet performance.

During each iteration of the downsampling pathway, a padded 3x3 covolutional layer is applied, followed by feeding layers through a batch normalization layer and a Rectified Linear Unit (ReLU) activation function. This process is repeated three times before downsampling with a 2x2 average pooling layer with a stride of 1. Average pooling layers are used instead of max pooling layers, as all image details and features are important to our objective. When all downsampling is completed, the network forms two branches that upscale separately, processing J_x and J_y, respectively. The network first performs a 2x2 upsampling convolution, then uses a skip connection to concatenate the channels from the downsampling pathway to the channels from the upsampling pathway of the same resolution. The process of applying a 3x3 covolutional layer, batch normalization, and ReLu activation occurs three times. Specific inter-layer connections and data resolutions at each step are depicted in Figure 1b.

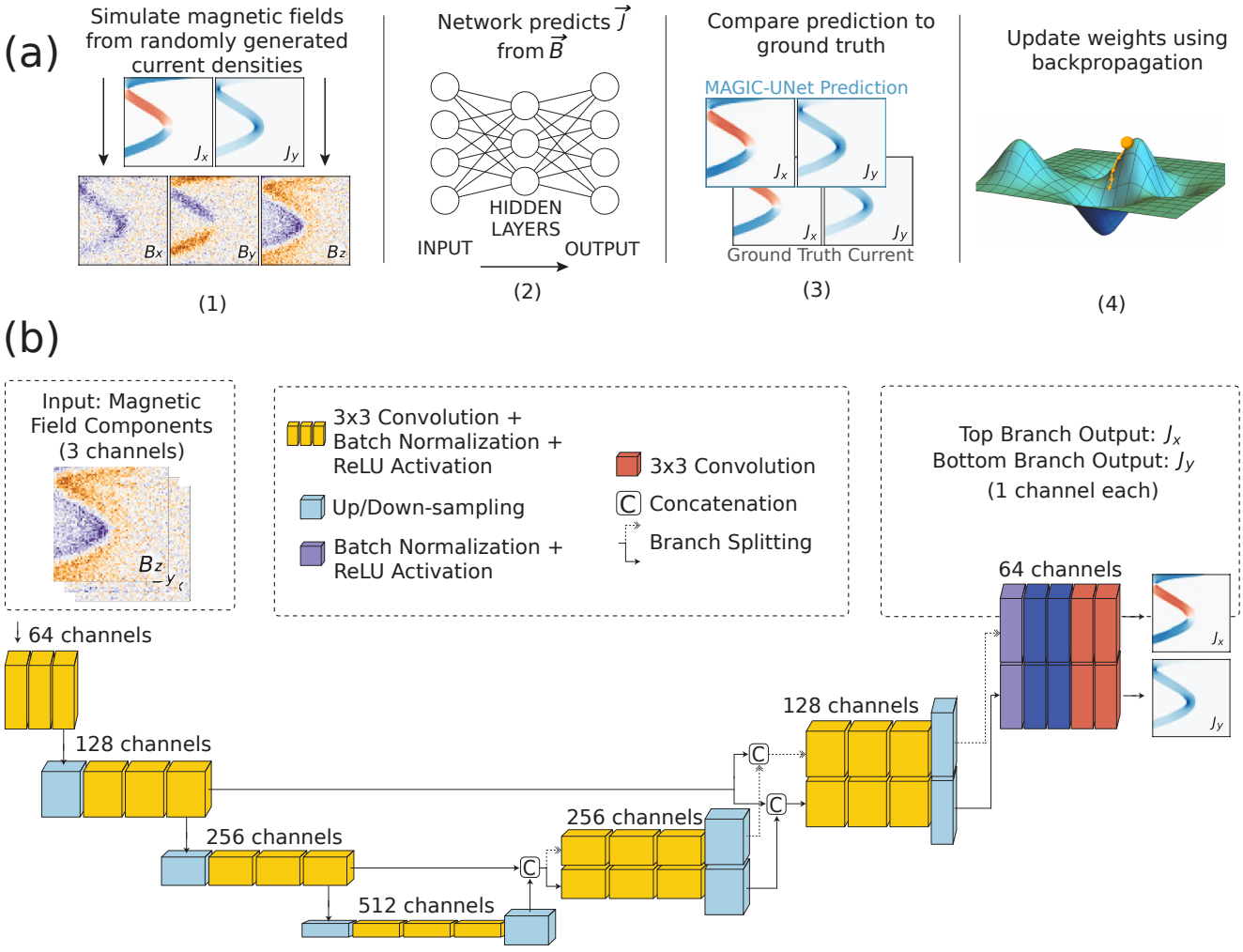

FIG. 1. (a) Illustration of the training process. The MAGIC-UNet network is trained on field-current pairs, taking inputs of vector magnetic field images and making predictions of planar current density distributions. After each batch, the predicted currents are compared to the ground truth using mean squared error; and then network weights are updated to improve the accuracy of the prediction. After many iterations the network will have learned to make quality predictions for unknown currents. (b) Diagram showing the structure of the MAGIC-UNet network. Each colored block represents a layer of the given number of channels, which is indicated above the blocks. Downsampling uses a 2x2 average pooling layer; upsampling also includes a 3x3 convolutional layer. Concatenation is represented by a C symbol and allows skip connections to join inputs from earlier parts of the network to later stages. The network has a branched structure with separate paths for J_x and J_y, which do not interact.

III. RESULTS

Using diverse validation datasets, the MAGIC-UNet network and Fourier Method are tasked with producing $J_{x,y}$ field inversions from vector magnetic field images both synthetic (Figure 2a) and experimental; and the performance of the two methods is compared. See Tables I and II, as well as discussion in the Supplemental Information. The accuracy of the inversions is evaluated by comparing to ground truth current densities using both qualitative and quantitative methods. As discussed below, MAGIC-UNet outperforms the Fourier Method by all metrics across all types of magnetic field input data.

We begin by assessing the performance of MAGIC-UNet and the Fourier Method on a synthetic validation data set, with standoff distances for each synthetic image sampled from a random distribution with mean 50 μm and standard deviation 10 μm to mimic the variation and uncertainty of an experimental dataset. While the exact standoff distance is used in the Fourier Method calculation, MAGIC-UNet is able to reconstruct accurate current densities without access to this information; a practical feature for experimental applications.

Based on the visual fidelity of current density maps, MAGIC-UNet predictions closely resemble ground truth J_x and J_y. The network performs notably well in addressing complex $\vec{J}$ features, including overlapping wire structures and sections with nonuniform current density

	MAGIC-UNet J_x	MAGIC-UNet J_y	Fourier Method J_x	Fourier Method J_y
Right Angle	0.895	0.898	0.421	0.421
All Angles	0.901	0.904	0.455	0.455
Thick Straight	0.900	0.902	0.359	0.359
Thin Curves	0.962	0.962	0.356	0.358
Thick Curves	0.950	0.947	0.396	0.399
Out-of-Distribution	0.948	0.947	0.519	0.514

TABLE I. Average Structural Similarity Index Measure (SSIM) for the J_x and J_y channels of MAGIC-UNet and the Fourier Method for the five classes of data in the validation set, as well as a set of out-of-distribution data. The Right Angle, All Angles, and Thick Straight classes can contain 1, 2, or 3 closely spaced current planes, while the Thin Curves and Thick Curves form shorts where currents overlap. Uncertainties for all average SSIM values are very small (≈ 0.0001).

	MAGIC-UNet J_x	MAGIC-UNet J_y	Fourier Method J_x	Fourier Method J_y
One Plane	0.966	0.969	0.411	0.412
Two Planes	0.902	0.907	0.415	0.414
Three Planes	0.842	0.845	0.406	0.403

TABLE II. Average SSIM for the J_x and J_y channels of MAGIC-UNet and the Fourier Method, for the three data classes that contain multiple closely spaced current planes. Uncertainties for all average SSIM values are very small ($\approx$0.0001).

(e.g., see Figure 2b). Comparatively, the Fourier Method consistently produces diffuse noise artifacts, blurring, and substantial smoothing of current edges. In addition to visual features, the peak current density values predicted by MAGIC-UNet are more accurate than those of the Fourier Method, which predicts peak current densities 2-3 times smaller than the true values; an example of this behavior is shown in Figure 2c.

For quantitative evaluations of inversion accuracy, we primarily employ the Structural Similarity Index Measure (SSIM), a metric used in image processing applications to examine preservation of structural information. SSIM compares the inversion to its ground truth. SSIM values of 0.95-0.99 are generally considered to have minor to imperceptible deviance from the true current density; values less than 0.88 are "poor" quality; and values below 0.50 are "bad" quality [53, 54]. Additionally, we expand our quantitative evaluation of inversion accuracy using Root Mean Square Error (RMSE) and Peak Signal to Noise Ratio (PSNR), which emphasize different traits of images.

To study the behavior of the MAGIC-UNet network and Fourier Method under variable levels of noise, we introduce artificial noise to the simulated input $\vec{B}$ data by applying an additive filter that adjusts pixel values by random amounts sampled from a Gaussian distribution. In our analysis, we define the noise level σ_n of a magnetic field image as the ratio of the magnetic field noise standard deviation σ to the maximum value of the magnetic field magnitude $\max(|\vec{B}|)$, across all image pixels. This normalization controls the Normal Distribution $\mathcal{N}(\mu = 0, \sigma_n^2)$ of the additive filter applied to the image. For example, a noise level of $\sigma_n = 0.2$ corresponds to setting the standard deviation σ of the distribution $\mathcal{N}$ used in the additive filter process to one-fifth of the maximum magnetic field magnitude in the image. In general, we find accurate predictions from MAGIC-UNet for $\sigma_n = 0.5$, roughly corresponding to an input image data SNR of 2. MAGIC-UNet is typically able to identify current structures for noise levels as large as $\sigma_n = 1.0$ (input SNR $\sim$ 1). However, network predictions are increasingly distorted as the noise level is increased; examples are shown in Figure 3a,b.

A. Quantitative Performance

The quantitative performance of MAGIC-UNet supports its visual fidelity. When inverting validation $\vec{B}$ data and comparing to the ground truth current density $\vec{J}_{GT}$, MAGIC-UNet achieves SSIM values above 0.95 (corresponding to good image quality) when given validation data with noise levels σ_n between 0 and 0.35 (input SNR > 2.8) ; and SSIM values above 0.88 (corresponding to "fair" image quality) for σ_n between 0.40 and 0.70 ($2.8 >$ input SNR > 1.4); see Figure 3c). Overall, SSIM values range from 0.99 at $\sigma_n = 0$ to 0.84 at $\sigma_n = 1$. In contrast, SSIM values for Fourier Method predictions span 0.71 to 0.32 over the same range, classifying as "poor" between noise levels of $\sigma_n = 0.0$ to $\sigma_n = 0.4$ (input SNR > 2.5) , and as "bad" for $\sigma_n > 0.45$. Throughout these tests, the MAGIC-UNet network does not favor any particular current density component: the difference in average SSIM for J_x and J_y does not exceed 1% of their total values, performing similarly in this regard to the Fourier inversions (see Tables I and II).

SSIM performance is similar across all types of in-distribution data, with the exception that distributions with multiple current planes lead to significantly worse average SSIM values for MAGIC-UNet predictions (Table II). To check for overfitting, we compare these results to the SSIM of reconstructions of out-of-distribution valida-

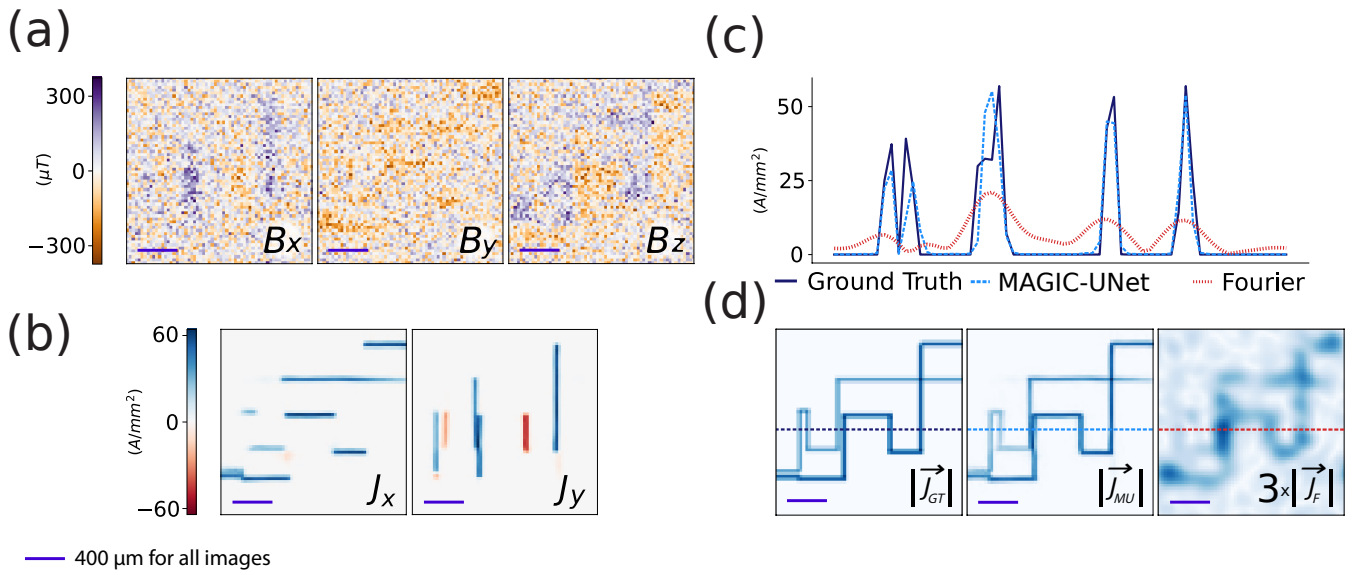

FIG. 2. Current density predictions for synthetic (simulated) vector magnetic field images. The validation (input) data set includes magnetic field images in a plane separated by 50 ± 10 μm from the planar current density distribution. (a) x, y, and z components of an example input magnetic field image, including added Gaussian noise at σ=0.5. (b) MAGIC-UNet current density predictions for input shown in (a), with x and y components shown separately. (c) Linecuts showing the current density magnitude comparing the ground truth (black), MAGIC-UNet (blue), and Fourier Method (red). The location of the linecuts is indicated by the dashed lines in (d). (d) Current density magnitude showing (left to right) the ground truth, MAGIC-UNet, and Fourier Method predictions for input shown in (a). The Fourier Method image has been globally multiplied by 3 to allow it to be visible on the same scale as the other two images. Scale bars are 400 μm for all planar images shown in (a), (b), and (d).

tion data. This out-of-distribution data, primarily consisting of straight and curved wire segments, is generated using a separate algorithm than for training and validation data, with corresponding magnetic fields simulated using COMSOL Multiphysics® instead of the analytic process described in Section V. (Also see discussion in the Supplemental Information.) For the example out-of-distribution data, MAGIC-UNet achieves SSIM of 0.962 $\sigma_n = 0.5$, which is within the range of the in-distribution single-layer datasets. However, when tasked with performing inversions on test data far outside of the distribution, such as intricate, branching wires with non-uniform width, MAGIC-UNet achieves a SSIM value of 0.88 at noise level $\sigma_n = 0.5$, showing that the network has not learned a fully general version of the inversion and may be biased towards distributions more similar to the training data sets. Interestingly, the gap between performance on in- and out-of-distribution validation data disappears at the highest noise levels, suggesting these effects are more important for ensuring high fidelity predictions given mild to moderate noise than for approximating the general features in a high noise image.

MAGIC-UNet predictions also exhibit less average error as indicated by the RMSE and PSNR metrics. Mean Square Error (MSE) and the related RMSE are frequently used to evaluate regression predictions, as MSE is the maximum likelihood estimator for data with Gaussian noise, such as our experimental data, and served as the loss function during training [55]. The average RMSE of Fourier Method predictions is about 3-4x greater than the MAGIC-UNet RMSE at every level of noise tested (Figure 3d). PSNR is perhaps the most widely used metric in image and video processing and seeks to quantify the ratio between reconstruction erors and the true signal [56]. PSNR for MAGIC-UNet predictions ranges from 41.7 dB at $\sigma_n = 0$ to 28.7 dB at $\sigma_n = 1$. Comparatively, PSNR for Fourier Method predictions ranges from 28.0 dB to 24.5 dB over the same interval (Figure 3e). The superior RMSE and PSNR for MAGIC-UNet relative to the Fourier Method is manifested by both fewer noise artifacts and more accurate current density predictions.

When tasked with inverting experimental data collected using a QDM, MAGIC-UNet shows similar performance improvements over the Fourier Method as seen with synthetic data, including fewer noise artifacts and more accurate wire widths. As an example, the predicted wire widths in Figure 4 are visibly much more accurate in MAGIC-UNet reconstructions than in Fourier Method reconstructions, when compared with experimental QDM magnetic images and wire layout ground truth.

However, more noise artifacts appear in MAGIC-UNet reconstructions of experimentally collected images than in reconstructions of synthetic images with the same σ_n. This result is likely due to spatially-correlated noise in the experimental images due to the fitting process for NV Optically Detected Magnetic Resonance (ODMR) spec-

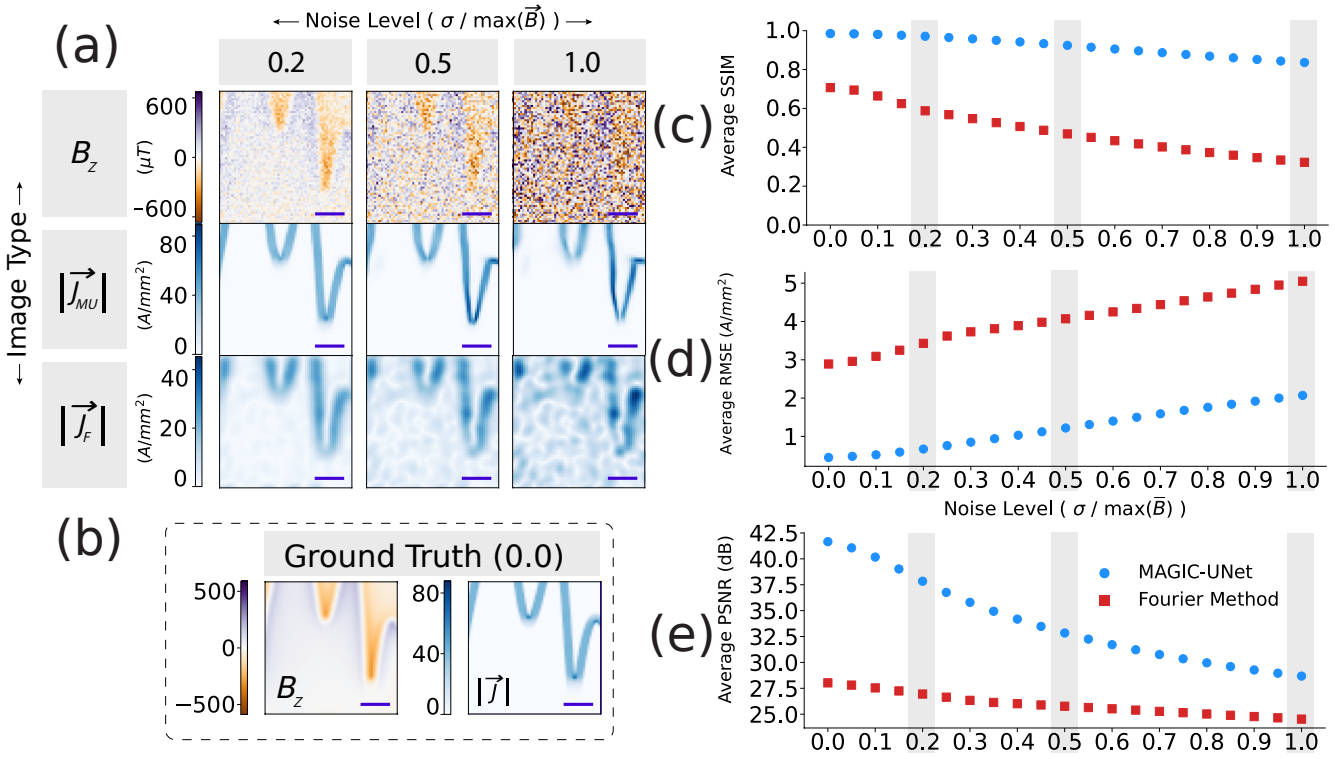

FIG. 3. Comparison of MAGIC-UNet and Fourier Method current density reconstructions for simulated input magnetic imaging data, with varying noise. The validation (input) data set includes magnetic field images in a plane separated by 40 ± 5 µm from the planar current density distribution. (a) z-component of an example input magnetic field image with added noise; and resulting magnitude of current density predictions for MAGIC-UNet and the Fourier Method. Three levels of noise are shown, for the same underlying ground truth, corresponding to $\sigma_n = 0.2$, 0.5, and 1.0. (b) Ground truth z-component of magnetic field image and current density magnitude. (c) Average SSIM for MAGIC-UNet (blue circles) and the Fourier Method (red squares) determined from validation data with different values of noise. (d) Average RMSE across the validation data for MAGIC-UNet and the Fourier Method. (e) Average PSNR across the validation data for MAGIC-UNet and the Fourier Method. Gray boxes in (c)-(e) correspond to noise levels displayed in (a). Scale bars are 400 µm for all planar images shown in (a) and (b).

tra; whereas the MAGIC-UNet network is trained on simulated magnetic field images with uncorrelated Gaussian noise. See discussion in the Supplemental Information.

B. Increasing Spatial Resolution

To facilitate inversions featuring spatially detailed current profiles, we assess two methods of increasing the input spatial resolution of the network. While the MAGIC-UNet model does not explicitly restrict the input size, the performance of the network degrades as the given input diverges from the image resolution on which the network is trained. Thus, we train a separate network using 256×256 data and a similar model, but with an additional downsampling and upsampling layer due to the larger input size. The resulting 256×256 resolution MAGIC-UNet performs well on inversions of synthetic data, outperforming its (nominal resolution) 64x64 counterpart for high σ_n and the 256×256 resolution Fourier Method at all σ_n (see Table III and Figure 5a,b). The 256×256 MAGIC-UNet network also performs well on low noise QDM experimental data, but degrades in performance compared to its 64x64 counterpart on higher-noise experimental data. This discrepancy may be explained by experimental image artifacts being significantly less Gaussian at 256×256 resolution than in a 64x64 pixel image, which is created through binning of the raw QDM experimental image.

Although the 256×256 MAGIC-UNet network performs well with the exception of high-noise experimental data, the enhanced resolution substantially increases the memory and computational time needed for training. To alleviate this shortcoming, we implement a tiling method that can effectively process larger images, with each image region employing the 64x64 resolution MAGIC-UNet network (Figure 5c). In this process, higher resolution images are segmented into overlapping 64x64 pixel tiles. MAGIC-UNet then processes each tile as if it is a separate data set. Once complete, overlapping tiles are trimmed and recombined into an image with the same resolution as the starting input. We test both 4 and 16 tile con-

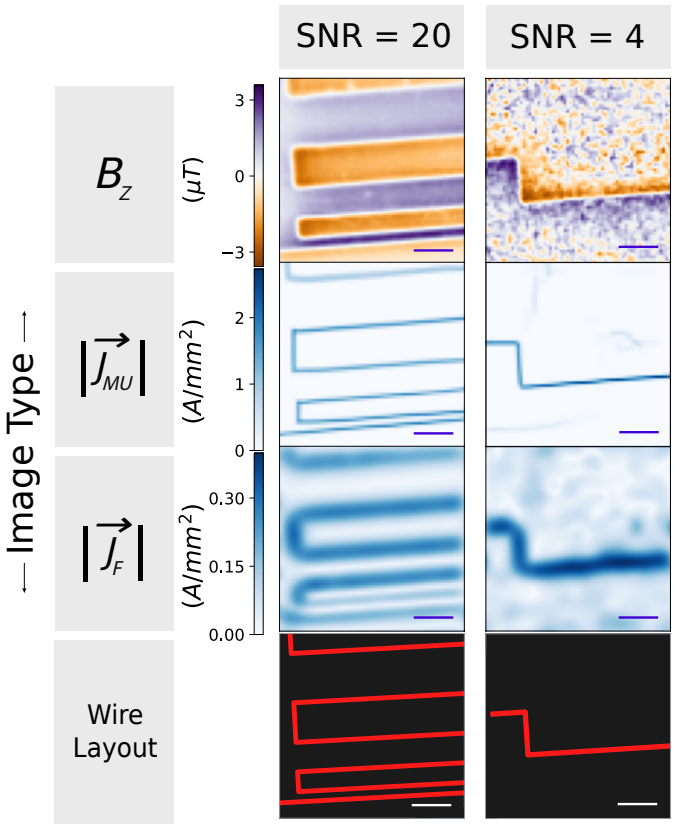

FIG. 4. Current inversion for two levels of SNR on experimental (QDM) magnetic field imaging data. The magnetic field images are in a plane separated from the current source by 65 μm (left column) and 36 μm (right column). Rows show the z-component of measured magnetic field images; the current density magnitudes from MAGIC-UNet and Fourier Method reconstructions; and the wire layout used to generate the source current for magnetic field images. Low SNR data is collected from a single-shot QDM image; while high SNR data is created from an average of 400 QDM images of the same magnetic field pattern. The wire layouts are cropped and rotated to match the QDM field of view. Scale bars are 400 µm for all planar images.

figurations on experimental data, allowing for resolution upscaling of approximately 2 and 4 times the linear size respectively.

As shown in Figure 5d, MAGIC-UNet makes satisfactory predictions using this tiling method, with superior SSIM to the full resolution Fourier Method for all noise levels. Note that the performance of the (tiled) MAGIC-UNet network is for a field of view up to four times smaller than the network training data, which is an interesting result to be studied in future work, given that the field of view significantly impacts the analytic transformation underlying the network architecture. The success of the tile predictions demonstrates that the MAGIC-UNet network can handle data outside of the parameter range on which it was trained. Further work could investigate whether training a network specifically to handle smaller fields of views for a tiling application leads to even better performance.

	SSIM $\sigma_n = 0.1$	SSIM $\sigma_n = 0.2$	SSIM $\sigma_n = 0.5$	SSIM $\sigma_n = 1.0$
64x64 MAGIC-UNet	0.981	0.971	0.925	0.836
256×256 MAGIC-UNet	0.976	0.975	0.958	0.923
256×256 Fourier Method	0.771	0.754	0.680	0.567

TABLE III. Average SSIM for 64x64 and 256×256 MAGIC-UNet and 256×256 Fourier Method for inversions of synthetic data with different σ_n.

C. Large Standoff

To examine the large standoff regime, we test a 64x64 (untiled) MAGIC-UNet network trained on magnetic field data simulated at a standoff distance of 500 ± 50 μm above the current plane, in contrast to the standoff distance of 50 ± 10 μm used in all other synthetic data in this work. This 500 μm distance corresponds to about 5 to 10 times the current source feature size and is therefore well beyond the regime where the Fourier Method is expected to perform well. Example results are shown in Figure 6a-d.

In this challenging regime, both reconstruction methods perform too poorly for our quantitative metrics to provide meaningful insight. However, it can be observed qualitatively that the MAGIC-UNet network performs better than the Fourier Method. For example, at $\sigma_n = 0.2$, Fourier Method predictions are severely hindered by input image blurring and edge effects, yielding spatially ill-defined current reconstructions in the general area of the current source. In comparison, MAGIC-UNet predicts spatially well-defined wires of roughly correct size, albeit with notably increased artifacts and inaccuracies compared to input data from a 50 μm standoff distance. Predictions on experimental data show a similar pattern: MAGIC-UNet makes more meaningful predictions for source current density than the Fourier Method due to its ability to yield current sources of reasonable width flowing in similar directions to the ground truth. Notable deviations of MAGIC-UNet from ground truth include rounded corners for wires, as visible in Figure 6e.

IV. DISCUSSION

The above results illustrate the utility of MAGIC-UNet for high-fidelity reconstruction of planar current density distributions, whether for synthetic or experimental data,

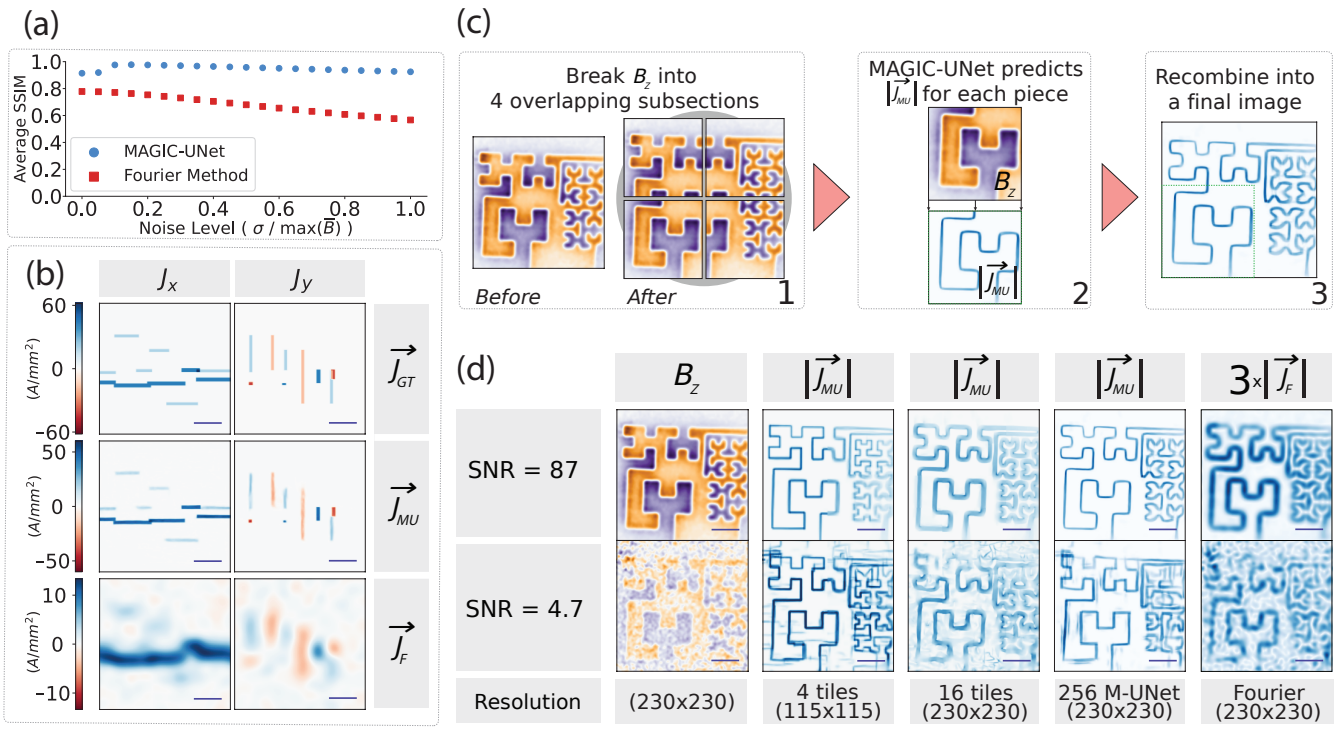

FIG. 5. Assessment of methods for efficient current density reconstruction of higher spatial resolution images than the (nominal) 64x64 resolution MAGIC-UNet network. (a) Comparison of average SSIM for 256×256 MAGIC-UNet and Fourier Method for synthetic data of varying σ_n. The synthetic magnetic field images are in a plane separated by 50 ± 10 µm from the planar current density distributions. (b) Comparison of current density reconstructions for (top to bottom) ground truth, 256×256 MAGIC-UNet, and Fourier Method on synthetic data with $\sigma_n = 0.5$ added noise. Columns show J_x and J_y. (c) Work flow of tiling process. (1) A magnetic field image is divided into overlapping 64x64 sections. (2) Each section is passed separately to the reconstruction protocol (MAGIC-UNet or Fourier Method). (3) The full reconstruction is stitched together from the individual predictions. Overlapping regions are divided equally between tiles, with duplicate predictions discarded. (d) Comparison of current density predictions using 4 different methods for both high and low SNR experimental data. QDM B_z images created by a Hilbert curve wire array are shown in column 1 with a resolution of 230x230 pixels. Current density reconstruction methods include a 4 tile implementation of 64x64 MAGIC-UNet producing 115x115 images (column 2); a 16 tile implementation of 64x64 MAGIC-UNet producing 230x230 images (column 3); a cropped version of (untiled) 256×256 MAGIC-UNet (column 4); and a 230x230 resolution version of the Fourier Method (column 5). The QDM magnetic field images are in a plane separated by 45 μm from the current source. Scale bars are 400 µm for all planar images shown in (b) and (d).

from two-dimensional images of vector Oersted magnetic fields. Performance of the MAGIC-UNet network, both qualitatively and quantitatively, is consistently superior to that of a conventional Fourier reconstruction method, including for cases that degrade the fidelity of both approaches: e.g., reduced input data SNR (a regime that also reduces the performance of physically informed networks [47]). The improvement is quite significant; e.g. applied to 50 μm standoff data, the 64x64 network performs about 10× better than the Fourier Method by RMSE with no noise, and about 2.5× better at SNR = 1. MAGIC-UNet performs better in SSIM, RMSE, and PSNR at $\sigma_n = 1$ than the Fourier Method does at $\sigma_n = 0$; showing the capacity to dramatically reduce measurement time which would otherwise be required to produce high SNR images for Fourier Method reconstruction. Using SNR=1 inputs for MAGIC-UNet reconstruction, which we found to be more accurate than input-SNR=20 Fourier Method reconstructions, would allow data to be collected 400 times faster.

With the ability to utilize faster experimental readout and associated weaker input signals, we are optimistic MAGIC-UNet could be a preferred alternative to the Fourier Method for current density reconstruction from magnetic imaging data.

A key challenge for future work will be assessing the utility of the MAGIC-UNet architecture for reconstruction of three-dimensional (3D) current sources from vector magnetic field images. Though the unconstrained (general) magnetic inverse problem does not have a unique solution, there are 3D solutions for specific cases of interest, e.g., when the current is constrained to discrete wires [57]. An intermediate step that may first be investigated is a circuit with current constrained to multiple planes plus occasional wires between planes. Current reconstruction of such a "2.5D" circuit might first

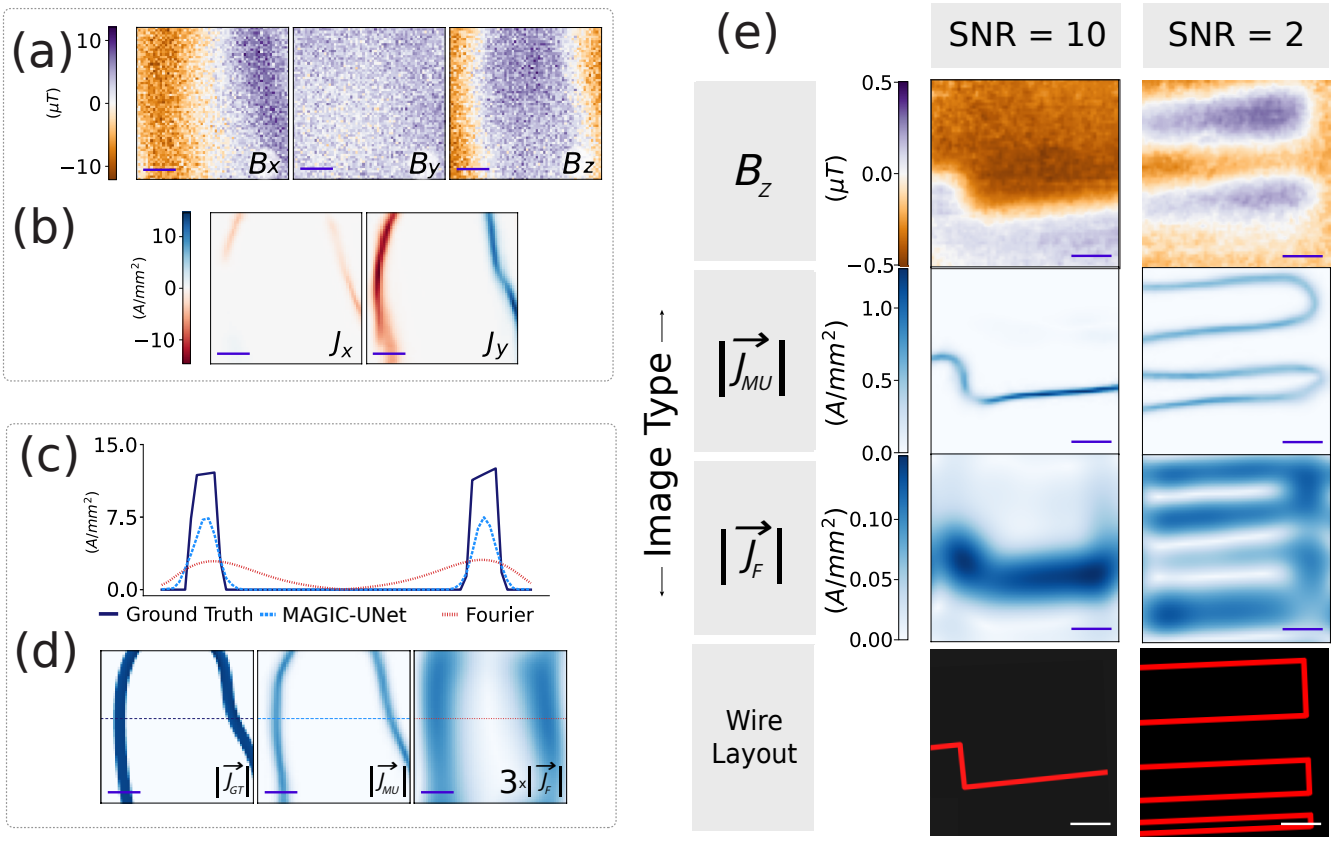

FIG. 6. Assessment of current density reconstruction methods at large standoff distances. (a) Images of x, y, and z components of magnetic field for synthetic data at 500 μm standoff with $\sigma_n = 0.2$. (c) J_x and J_y predictions from MAGIC-UNet for input data in (a). (c) Linecuts comparing current density for ground truth (black), MAGIC-UNet (blue), and Fourier Method (red). (d) Comparison of ground truth current density magnitude image with MAGIC-UNet and Fourier Method predictions. Fourier Method prediction is shown at 3x magnitude in order for it to be visible in the same data range as the other two plots. Dashed lines show the location of linecuts plotted in (c). (e) QDM B_z images, MAGIC-UNet $|\vec{J}|$, and Fourier Method $|\vec{J}|$ for high-SNR and low-SNR experimental data taken at standoff distance $\approx$ 500 μm. Scale bars are 400 µm for all planar images shown.

separate magnetic field images associated with distinct planes before determining the associated current densities. $\vec{J}$ data could be handled similarly to the approach of MAGIC-UNet through an iterative process removing the topmost layer, or differently as a single step that performs the entire transformation for the entire multiple-plane circuit. Finally, another challenge for future work is to adapt the MAGIC-UNet methodology to magnetization distributions or combined current and magnetization distributions.

V. METHODS

A. Synthetic Data Creation and Training

We developed a data generation workflow to create 153,600 synthetic $\vec{J}$-$\vec{B}$ current-field pairs for model training and 4096 field-current pairs for validation and testing. Starting with random shapes produced by multiple wire-generation algorithms, finite element simulations in both COMSOL Mutliphysics® and MATLAB calculate the planar current density resulting from a voltage difference across the boundary of the simulated conductive material. To ensure robust training, we include a wide variety of wires, including shorts, closely spaced current layers, and non-uniform current density around corners. Once the current density simulations are complete, Oersted magnetic fields are calculated analytically using the Biot-Savart law. Additional details regarding training data generation are provided in the Supplementary Information.

The RMSProp optimization algorithm [58] is employed for training the network. Training details are documented in Table IV. The approach to data creation and training does not require access to specialized equipment or extensive computational time, allowing models to be trained on demand for different needs. The entire process from creating the training data to training the model requires about 24 hours. Once the network is trained, current density predictions can be generated for thousands

of magnetic field images in a few seconds, making the method extremely efficient.

Learning Rate	Epochs	Mini-batch	Loss Function	Library	Graphics Processor
0.0001	50	32	MSE	TensorFlow 2.10.0	NVidia RTX-A6000

TABLE IV. MAGIC-UNet training details using RMSProp optimization algorithm. The learning rate is set to 0.0001, while the training process utilizes 50 epochs with 32 image minibatch size. The loss function is designed to minimize the average mean squared error between the ground truth current densities and output images of the inference process. Total training time is 11 hours using the TensorFlow 2.10.0 machine learning framework [59] and an rtx-a6000 graphics processing unit with 48 GB memory capacity.

B. Experimental Data

To investigate the utility of the MAGIC-UNet for experimental applications, we used QDM magnetic imaging data, both from a recently published study [9] (by our group and collaborators) of a custom printed circuit board (PCB); and new data, acquired with the same QDM and PCB at variable standoff distances and noise levels. QDM magnetic images are obtained using a 4×4 mm ensemble NV diamond with a 1.7 μm NV surface layer. (Details about the QDM are described in [9].) A 5 mA current is applied to the custom PCB to produce $\tilde{5}\,\mu$T Oersted magnetic fields measured by the NV centers. Magnetic field images are collected for three different current geometries, with standoff distances ranging from 36 to 65 μm for small standoff and 486 to 650 μm for large standoff. Larger standoff distances employ a translation stage to separate the diamond from the PCB. Both single-shot and averaged measurements (20 or 400 exposures) are obtained, generating data with SNR between 4 and 87 for small standoff, and between 2 and 10 for large standoff. The resolution of the experiment is limited by resolution of the camera, with a pixel size of about 2 μm in the raw data; after cropping and binning to a 64x64 image each pixel is approximately 31 μm. See the Supplemental Information for more information about the experimental process.

C. Fourier Method Optimization

Fourier space transformations described by [31], [60] are implemented to produce synthetic magnetic field data B_x, B_y and B_z from randomly generated, planar-current densities J_x and J_y at variable standoff distances. These standoff distances form a normal distribution with mean 50 μm and standard deviation of 10 μm, which approximates the experimental range of separation between the NV sensing layer and current traces in the PCB (see discussion in the Supplemental Information).

In order to show the utility of the MAGIC-UNet method, we implement careful optimization of the Fourier Method. B_x and B_y field components are used rather than B_z for a more robust inversion technique, as described in [60]. In addition, Gaussian filter processing on magnetic field images is necessary to prevent errors from aliasing effects and a finite Fourier bandwidth. A Hann window function improves the resilience against noise of the Fourier inversion [38]. With these filters implemented, a 3D parameter optimization search is performed to find optimal Hann window and Gaussian filter strengths for multiple standoff distances, maximizing SSIM for synthetic data with respect to ground truth current densities.

VI. DATA AVAILABILITY

Trained models, validation and experimental test data, and associated scripts are available at: https://doi.org/10.7910/DVN/SD6PVP [61]. Full size datasets for training additional models can be found at https://doi.org/10.7910/DVN/SJDS2O and https://doi.org/10.7910/DVN/QPCSOI.

VII. ACKNOWLEDGEMENTS

We thank Bo Zhu and Edlyn Levine for their helpful contributions to the early development of this project; and Stephen DeVience for suggesting the acronym MAGIC-UNet. M.S.R. acknowledges the generous support of the Kiyomi and Ed Baird MGH Research Scholar award. This material is based on work by N.R.R. and D.B. supported by the National Science Foundation Graduate Research Fellowship Program under Grant Nos. DGE 2236417 and DGE 1840990. R.L.W. acknowledges support from the Laboratory for Physical Sciences Jumping Electron Quantum Fellowship Program under Award No. H9823022C0029; the IC Postdoctoral Fellowship under Award No. STEMWD00851; the U.S. Air Force Office of Scientific Research under Grant No. FA9550-22-1-0312, and the University of Maryland Quantum Technology Center. S.M.O. acknowledges support from the MITRE Independent Research and Development Program. K. S. O. acknowledges support through an appointment to the Intelligence Community Postdoctoral Research Fellowship Program at the University of Maryland, administered by Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Office of the Director of National Intelligence. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies.

Approved for Public Release; Distribution Unlimited. Public Release Case Number 24-1979. Portions of this technical data were produced for the U. S. Government under Contract No. FA8702-19-C-0001 and W56KGU-18-D-0004, and is subject to the Rights in Technical Data-Noncommercial Items Clause DFARS 252.227-7013 (FEB 2014). ©2024 The MITRE Corporation.

VIII. AUTHOR CONTRIBUTION STATEMENT

*Denotes equal contribution between N.R.R. and D.B..

M.J.T. conceived the project and developed the early ideas with M.J.H.K. and D.B.. D.B. designed the MAGIC-UNet architecture and performed network optimization. D.B., N.R.R., and N.L.. worked on the machine learning code base and optimization. N.R.R. performed data generation and network training, optimized hyperparameters and characteristics of training data, managed and modified the project's central code base, coordinated other researchers on the creation of experimental and synthetic data, and developed and carried out the data analysis. S.M.O. and M.J.T. built the experimental setups and collected the experimental data, with assistance from K.S.O.. M.J.T., N.R.R., J.T., and D.M.D. created the synthetic data used for training and validation. D.M.D. implemented the Fourier Method analysis and optimized parameters for its performance. N.R.R. and D.M.D. wrote the manuscript, created data visualization tools, and generated the figures. R.L.W., M.S.R., M.J.H.K., and K.S.O. supervised the project. All authors reviewed the manuscript and were able to contribute to its editing and final content.

[1] A. Orozco, EDFA Technical Articles (2009).

[2] A. Nowodzinski, M. Chipaux, L. Toraille, V. Jacques, J.-F. Roch, and T. Debuisschert, Microelectronics Reliability **55**, 1549 (2015).

[3] F. Infante, P. Perdu, S. Petremont, and D. Lewis, in *2009 16th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits* (2009) pp. 208–212, iSSN: 1946-1550.

[4] *Backside Integrated Circuit Magnetic Field Imaging with a Quantum Diamond Microscope*, International Symposium for Testing and Failure Analysis (ASM International, 2020).

[5] P. Kehayias, E. V. Levine, L. Basso, J. Henshaw, M. Saleh Ziabari, M. Titze, R. Haltli, J. Okoro, D. R. Tibbetts, D. M. Udoni, E. Bielejec, M. P. Lilly, T.-M. Lu, P. D. D. Schwindt, and A. M. Mounce, Physical Review Applied **17**, 014021 (2022).

[6] M. J. Turner, *Quantum Diamond Microscopes for Biological Systems and Integrated Circuits*, Ph.D. thesis, Harvard University (2020).

[7] M. J. Turner, N. Langellier, R. Bainbridge, D. Walters, S. Meesala, T. M. Babinec, P. Kehayias, A. Yacoby, E. Hu, M. Lončar, *et al.*, Physical Review Applied **14**, 014097 (2020).

[8] F. Infante, P. Perdu, and D. Lewis, Microelectronics Reliability 20th European Symposium on the Reliability of Electron Devices, Failure Physics and Analysis, **49**, 1169 (2009).

[9] *47th International Symposium for Testing and Failure Analysis*, International Symposium for Testing and Failure Analysis (ASM, 2021).

[10] *Construction of a 3-D Current Path Using Magnetic Current Imaging*, 33rd International Symposium for Testing and Failure Analysis (ASM International, 2007).

[11] Y. Hu, G. Z. Iwata, M. Mohammadi, E. V. Silletta, A. Wickenbrock, J. W. Blanchard, D. Budker, and A. Jerschow, Battery Diagnostics with Sensitive Magnetometry (2019), arXiv:1905.12507 [physics].

[12] Y. Hu, G. Z. Iwata, M. Mohammadi, E. V. Silletta, A. Wickenbrock, J. W. Blanchard, D. Budker, and A. Jerschow, Proceedings of the National Academy of Sciences **117**, 10667 (2020), publisher: Proceedings of the National Academy of Sciences.

[13] M. G. Bason, T. Coussens, M. Withers, C. Abel, G. Kendall, and P. Krüger, Journal of Power Sources **533**, 231312 (2022).

[14] F. Brauchle, F. Grimsmann, O. von Kessel, and K. P. Birke, Journal of Power Sources **507**, 230292 (2021).

[15] S. Scholten, G. Abrahams, B. Johnson, A. Healey, I. Robertson, D. Simpson, A. Stacey, S. Onoda, T. Ohshima, T. Kho, J. Ibarra Michel, J. Bullock, L. Hollenberg, and J.-P. Tetienne, Physical Review Applied **18**, 014041 (2022), publisher: American Physical Society.

[16] K. Kaufmann, D. Lausch, C.-M. Lin, M. Rudolph, D. Hahn, and M. Patzold, physica status solidi (a) **218**, 2000292 (2021).

[17] N. Chikumoto, S. Yamaguchi, and O. Shyshkin, IEEE Transactions on Applied Superconductivity **31**, 1 (2021), conference Name: IEEE Transactions on Applied Superconductivity.

[18] S. Sakai, Y. Yamada, Y. Zhao, J. Zhu, and M. Inoue, IEEE Transactions on Applied Superconductivity **33**, 1 (2023), conference Name: IEEE Transactions on Applied Superconductivity.

[19] P. Bevington, R. Gartman, W. Chalupczak, C. Deans, L. Marmugi, and F. Renzoni, Applied Physics Letters **113**, 063503 (2018).

[20] J. P. Wikswo, in *SQUID Sensors: Fundamentals, Fabrication and Applications*, NATO ASI Series, edited by H. Weinstock (Springer Netherlands, Dordrecht, 1996) pp. 629–695.

[21] Z. Wu, K. Higashikawa, K. Imamura, Z. Xu, Y. Ma, and T. Kiss, IEEE Transactions on Applied Superconductivity **31**, 1 (2021), conference Name: IEEE Transactions on Applied Superconductivity.

[22] J. R. Kirtley, Reports on Progress in Physics **73**, 126501 (2010).

[23] E. Marchiori, L. Ceccarelli, N. Rossi, G. Romagnoli, J. Herrmann, J.-C. Besse, S. Krinner, A. Wallraff, and M. Poggio, Applied Physics Letters **121**, 052601 (2022).

[24] M. J. H. Ku, T. X. Zhou, Q. Li, Y. J. Shin, J. K. Shi, C. Burch, L. E. Anderson, A. T. Pierce, Y. Xie, A. Hamo, U. Vool, H. Zhang, F. Casola, T. Taniguchi, K. Watanabe, M. M. Fogler, P. Kim, A. Yacoby, and R. L. Walsworth, Nature **583**, 537 (2020).

[25] Y. Feng, J. Zhu, W. Lin, Z. Lian, Y. Wang, H. Li, H. Yao, Q. He, Y. Pan, Y. Wu, J. Zhang, Y. Wang, X. Zhou, J. Shen, and Y. Wang, Nano Letters **22**, 7606 (2022), publisher: American Chemical Society.

[26] M. Palm, W. Huxter, P. Welter, S. Ernst, P. Scheidegger, S. Diesch, K. Chang, P. Rickhaus, T. Taniguchi, K. Watanabe, K. Ensslin, and C. Degen, Physical Review Applied **17**, 054008 (2022), publisher: American Physical Society.

[27] A. V. Bjørlig, D. V. Christensen, R. Erlandsen, N. Pryds, and B. Kalisky, ACS Applied Electronic Materials **4**, 3421 (2022), publisher: American Chemical Society.

[28] L. Basso, P. Kehayias, J. Henshaw, M. Saleh Ziabari, H. Byeon, M. P. Lilly, E. Bussmann, D. M. Campbell, S. Misra, and A. M. Mounce, Nanotechnology **34**, 015001 (2023).

[29] S. E. Lillie, N. Dontschuk, D. A. Broadway, D. L. Creedon, L. C. Hollenberg, and J.-P. Tetienne, Physical Review Applied **12**, 024018 (2019), publisher: American Physical Society.

[30] A. Wickenbrock, N. Leefer, J. W. Blanchard, and D. Budker, Applied Physics Letters **108**, 183507 (2016).

[31] B. J. Roth, Sensors **23**, 4218 (2023).

[32] U. Leder, H.-P. Pohl, S. Michaelsen, T. Fritschi, M. Huck, J. Eichhorn, S. Müller, and H. Nowak, International Journal of Cardiology **64**, 83 (1998).

[33] R. Merwa, K. Hollaus, P. Brunner, and H. Scharfetter, Physiological Measurement **26**, S241 (2005).

[34] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, Reviews of Modern Physics **65**, 413 (1993), publisher: American Physical Society.

[35] R. Alvarez, IEEE Transactions on Medical Imaging **9**, 299 (1990), conference Name: IEEE Transactions on Medical Imaging.

[36] S. Knappe, O. Alem, D. Sheng, and J. Kitching, Journal of Physics: Conference Series **723**, 012055 (2016), publisher: IOP Publishing.

[37] M. Limes, E. Foley, T. Kornack, S. Caliga, S. McBride, A. Braun, W. Lee, V. Lucivero, and M. Romalis, Physical Review Applied **14**, 011002 (2020), publisher: American Physical Society.

[38] B. J. Roth, N. G. Sepulveda, and J. P. Wikswo, Journal of Applied Physics **65**, 361 (1989).

[39] E. V. Levine, M. J. Turner, P. Kehayias, C. A. Hart, N. Langellier, R. Trubko, D. R. Glenn, R. R. Fu, and R. L. Walsworth, Nanophotonics **8**, 1945 (2019).

[40] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, IEEE Transactions on Image Processing **26**, 4509 (2017), arXiv: 1611.03679.

[41] G. Wang, J. C. Ye, and B. De Man, Nature Machine Intelligence **2**, 737 (2020).

[42] J. Schlemper, I. Oksuz, J. R. Clough, J. Duan, A. P. King, J. A. Schnabel, J. V. Hajnal, and D. Rueckert, arXiv preprint arXiv:1909.10995 (2019).

[43] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, Nature **555**, 487 (2018), number: 7697 Publisher: Nature Publishing Group.

[44] E. Coccorese, R. Martone, and F. Morabito, IEEE Transactions on Magnetics **30**, 2829 (1994), conference Name: IEEE Transactions on Magnetics.

[45] M. Kishimoto, K. Sakasai, and K. Ara, Journal of Applied Physics **79**, 1 (1996).

[46] M. T. McCann, K. H. Jin, and M. Unser, IEEE Signal Processing Magazine **34**, 85 (2017), conference Name: IEEE Signal Processing Magazine.

[47] A. Dubois, D. Broadway, A. Stark, M. Tschudin, A. Healey, S. Huber, J.-P. Tetienne, E. Greplova, and P. Maletinsky, Untrained physically informed neural network for image reconstruction of magnetic field sources (2022).

[48] M. A. Tschudin, D. A. Broadway, P. Siegwolf, C. Schrader, E. J. Telford, B. Gross, J. Cox, A. E. E. Dubois, D. G. Chica, R. Rama-Eiroa, E. J. G. Santos, M. Poggio, M. E. Ziebel, C. R. Dean, X. Roy, and P. Maletinsky, Nature Communications **15**, 6005 (2024), publisher: Nature Publishing Group.

[49] O. Ronneberger, P. Fischer, and T. Brox, arXiv:1505.04597 [cs] (2015), arXiv: 1505.04597.

[50] D. Mehta, D. Padalia, K. Vora, and N. Mehendale, in *2022 5th International Conference on Advances in Science and Technology (ICAST)* (IEEE, 2022) pp. 306–313.

[51] H. Chen, Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou, and G. Wang, Biomedical Optics Express **8**, 679 (2017).

[52] Y. Bengio, P. Simard, and P. Frasconi, IEEE transactions on neural networks **5**, 157 (1994).

[53] M. Zanforlin, D. Munaretto, A. Zanella, and M. Zorzi, in *2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)* (IEEE, Hammamet, Tunisia, 2014) pp. 656–661.

[54] J. R. Flynn, S. Ward, J. Abich, and D. Poole, in *Engineering Psychology and Cognitive Ergonomics. Understanding Human Cognition*, Lecture Notes in Computer Science, edited by D. Harris (Springer, Berlin, Heidelberg, 2013) pp. 23–30.

[55] T. O. Hodson, Geoscientific Model Development **15**, 5481 (2022).

[56] S. Winkler and P. Mohandas, IEEE Transactions on Broadcasting **54**, 660 (2008).

[57] K.-H. Hauer, L. Kühn, and R. Potthast, Inverse Problems **21**, 955 (2005).

[58] G. "Hinton, N. Srivastava, and K. Swersky, "rmsprop: Divide the gradient by a running average of its recent magnitude (neural networks for machine learning)." (2018).

[59] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, *et al.*, arXiv preprint arXiv:1603.04467 (2016).

[60] D. Broadway, S. Lillie, S. Scholten, D. Rohner, N. Dontschuk, P. Maletinsky, J.-P. Tetienne, and L. Hollenberg, Physical Review Applied **14**, 024076 (2020).

[61] N. Reed, D. Bhutto, M. Turner, D. Daly, S. Oliver, J. Tang, K. Olsson, M. Ku, M. Rosen, and R. Walsworth, Replication Data for: Machine Learning for Improved Current Density Reconstruction from 2D Vector Magnetic Images (2025).

SUPPLEMENTAL INFORMATION

A. Experimental Data

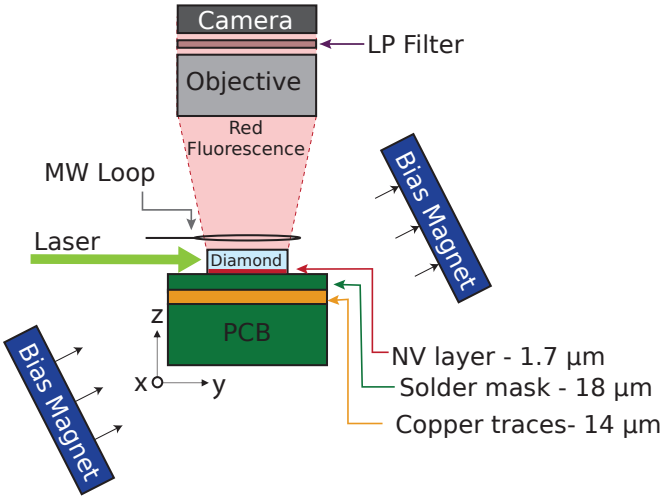

FIG. S1. Schematic of experimental set-up. Magnetic field imaging is performed by a quantum diamond microscope (QDM) employing Continuous Wave Optically-Detected Magnetic Resonance (CW-ODMR) measurements of the NV sensing layer at the surface of a diamond chip. The diamond is placed in close proximity to the custom PCB with the NV layer facing downwards. A wire loop above the diamond delivers a microwave (MW) signal for CW-ODMR measurements, driving the NV electronic spins. Permanent magnets are positioned to produce a static bias field $\approx 5\,\mathrm{mT}$ aligned to spectrally separate the four NV orientations. A green laser optically excites the NV ensemble, producing red photoluminescence focused by an optical objective. A long pass (LP) filter blocks unwanted green light from the CMOS Camera. Data is generated by running current through the copper trace layer of the custom PCB, producing magnetic fields imaged by the QDM. Precision stages (not shown) adjust the standoff separation of the planar currents in the PCB from the NV sensing layer.

We obtain two sets of QDM experimental data using a Continuous Wave Optically-Detected Magnetic Resonance (CW-ODMR) imaging technique: in one "signal" data set, the PCB is electrically active and producing magnetic fields; whereas for a second "reference" data set only the ambient magnetic field environment is recorded. We then fit signal and reference data with custom fitting software to extract vector magnetic field components at each pixel of the images, subtracting reference data from signal data to remove artifacts, e.g., from ferromagnetic components of the PCB and gradients in the bias magnetic field.

To create different levels of noise in the images, we vary the amount of averaging of the measured CW-ODMR data before the fit is performed to determine magnetic field values per pixel. For high SNR images, we average 400 measurements at each MW frequency of the CW-ODMR spectrum; whereas for the low SNR image we perform the fit on a single sweep of MW frequency with no signal averaging.

The QDM employs a pair of 2 inch SmCo magnets to produce a bias field $\approx 5\,\mathrm{mT}$, aligned such that the ODMR spectra of the 4 NV orientations are separated, allowing the reconstruction of vector magnetic field measurements. The diamond, fabricated by Element Six™, features a $1.7\,\mu\mathrm{m}$ NV layer enriched with nitrogen-15 ([N]=17 ppm, [NV]=2 ppm). The diamond is isotopically purified to contain ~ 99.995 % carbon-12.

For CW-ODMR measurements, a Stanford Research Systems SG384 signal generator produces a sinusoidal MW waveform resonate with the NV spin transitions. The signal is amplified by a Mini Circuits ZHL-16W-43-S+ high power amplifier and delivered through a custom made 7 mm diameter wire loop located above the diamond. The NV layer is optically excited with 1.5 W of power supplied by a Lighthouse Photonics Sprout-H-10W laser. A 4x optical objective (Olympus UPlanFL N 0.13 NA), a 633 nm longpass filter (Semrock LP02-633RU-25), and a CMOS CCD camera (Basler acA1920-155um) form the NV photoluminescence imaging system. A Rigol DG1022U arbitrary waveform generator produces a 100 mV DC signal used to drive currents through copper wire traces in the custom PCB, which is mounted on precision mechanical stages below the diamond. The stages allow adjustments of the standoff distance of the planar currents in the PCB from the NV sensing layer. Figure S1 shows a schematic of the experimental set-up.

B. Image SNR

To determine the signal-to-noise ratio (SNR) of experimental magnetic field images, we first find the maximum field strength along a line cut orthogonal to the direction of a wire segment producing a strong signal. Then, we compute the average signal by repeating this process at each pixel along the entire length of the segment. In some images, such as from the Hilbert curve shown in Figure 5d, the strongest wire segments produce fields 2-3 times greater than the weakest signal segments due to cancellation of magnetic fields near closely spaced wires. The average noise is determined by computing the standard deviation of pixels in a low-signal region of the image.

C. Noise Analysis

To inform the generation of synthetic data with artificial noise that properly reflects observations in experimental measurements, we analyze the noise characteristics of QDM magnetic field images, such as those shown in Figure 5. For numerous experimental datasets, we subtract high SNR images (constructed from the average of 400 CW-ODMR spectral sweeps with resulting

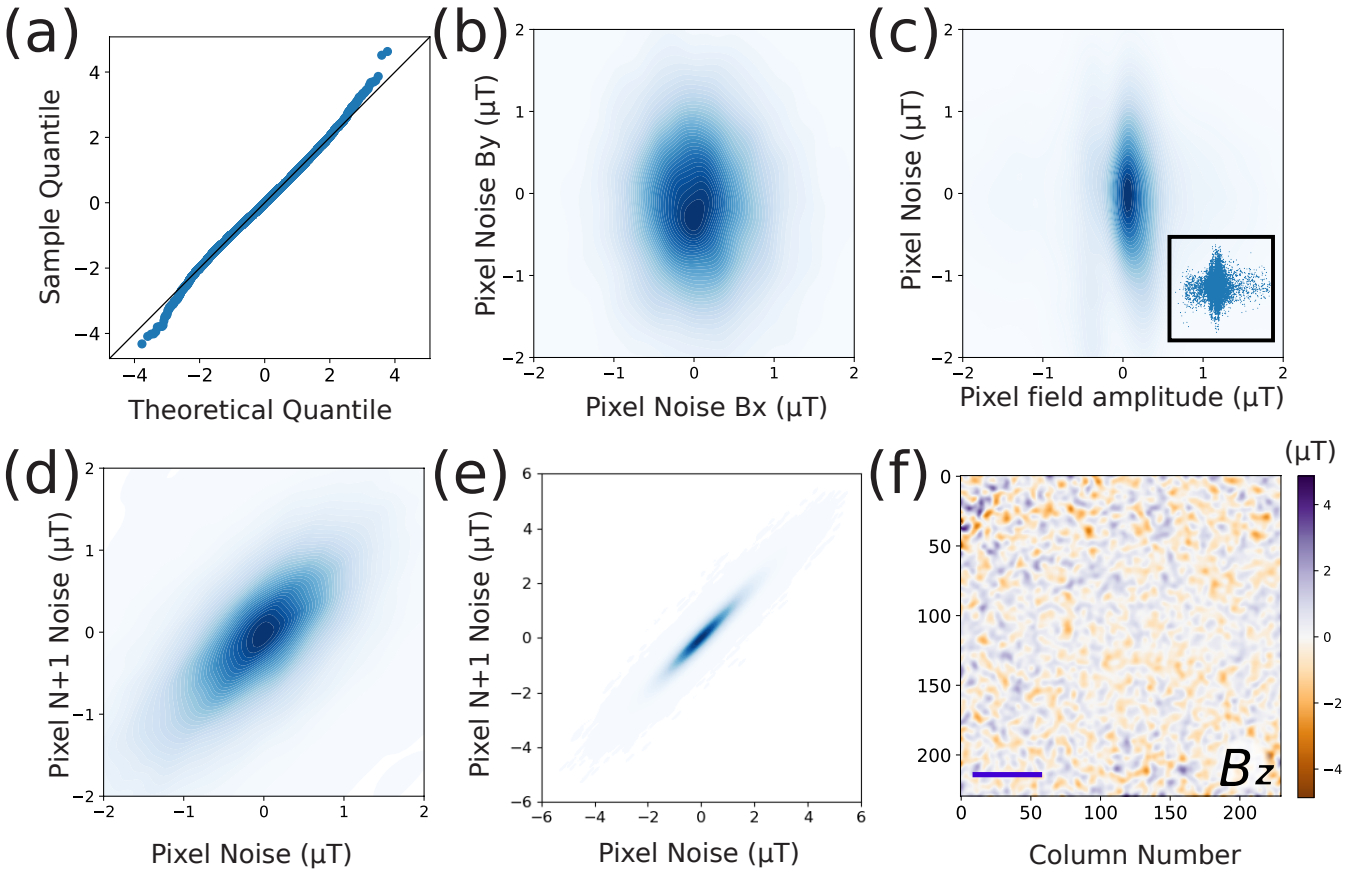

FIG. S2. Analysis of the noise data set produced from experimental QDM magnetic field images such as those shown in Figures 4 and 5. Seven pairs of low and high SNR images are used to generate the data set, and are compared at 230x230 resolution and at 64x64 resolution after binning and cropping. The noise data set is generated by subtracting high SNR magnetic field images from low SNR images with the same field of view. This noise data set retains vector information, thus consisting of x-noise, y-noise, z-noise images. (a) Quantile-quantile plot of the noise data set exhibits approximately linear dependence, indicating a normal distribution. (b) Per-pixel x-noise vs per-pixel y-noise of all image pixels. Noise is uncorrelated between vector components. (c) Per-pixel noise vs magnetic field amplitude (of the unprocessed, high-SNR images used to create the noise data set of all image pixels). There is minimal correlation between magnetic field amplitude and noise. Inset shows a scatter plot of this plot for a single image pair to better illustrate the behavior where magnetic field amplitude is large. (d) Noise data set is binned to 64x64 resolution; per pixel amplitude noise for row N, column M pixels plotted against row N+1, column M pixels. There is significant spatial correlation ($R^2 = 0.60$). (e) Same as (d), binned to 230x230 final resolution. There is larger spatial correlation ($R^2 = 0.95$). (f) Example B_z component contributing to the noise data set, generated by subtracting a high-SNR iamge from a low-SNR image. Noise artifacts $\approx$ 5-10 pixels can be seen. Scale bar is 400 μm for the planar B_z image shown in (f).

SNR $\approx$ 87) from their low-quality counterparts within the same field of view (a single CW-ODMR sweep with SNR $\approx$ 4.7) to produce noise datasets. Analysis of these datasets shows that the experimental noise is approximately Gaussian (see Figure S2a), with no significant correlation among field components (Figure S2b). This analysis also indicates that the noise is additive, since the noise distribution does not change as a function of field amplitude (Figure S2c).

Noise spatial correlations are visually apparent as cluster artifacts in QDM magnetic field images, approximately 5-10 pixels wide (Figure S2f); and manifest in Figure S2d,e as interdependent variables. This modest deviation from a purely additive, Gaussian model negatively impacts the performance of the MAGIC-UNet network when applied to experimental data, as the network is trained only on synthetic data utilizing uncorrelated Gaussian noise. For example, MAGIC-UNET performs less well for the 230x230 resolution experimental data set (spatial correlation $R^2 = 0.95$) than for the 64x64 resolution data (spatial correlation $R^2 = 0.60$), as seen in Figure S2d,e. In future work, these issues may be addressed by (i) reducing spatial correlations in experimental data, whether induced by the hardware or the fitting algorithm; and/or (ii) introducing spatial correlations into the synthetic training data.

D. Synthetic Data Generation

Synthetic data generation incorporates several key steps to ensure robust training of the MAGIC-UNet. To begin, we generate randomized planar current density images by varying total current, wire widths, numbers of bends, locations and angles of wires, and standoff distances across two classes of wire types. "Class I" features curving and looping wires with up to 20 bends, allowing wires to form shorts if they cross (see Figure S3). This class aids the network in learning about non-uniform current flow. "Class II" wires consist of only straight segments, changing directions at arbitrary angles and forming random intersections (see Figure S4). We add current densities together at these intersections to emulate the effect of separated wires crossing in chips featuring insulating layers. Key parameters used are listed in Table S1. We choose parameters both to approximate common behaviors in our PCB sample and support a robust and unbiased model.

After generating synthetic data, we increase image diversity by employing three separate unitary operations: 90 degree rotation, horizontal reflection, and vertical reflection. For a given image, each operation occurs with 50% probability, resulting in 8 possible permutations. This process helps eliminate sources of bias in current density components, directions, and spatial locations. Additionally, approximately 2% of the training data features wires exclusively outside of the field of view, in order to teach the network to understand fields generated outside the region of interest.

We then produce a final set of current densities using the Biot-Savart law recast as a Fourier Transform. This Forward Fourier Method produces a high fidelity solution since the current density images do not have noise added at this stage in the synthetic data generation process. We normalize each input magnetic field image individually to create a more optimal normalization scheme, as the input data has a wide spread of maximum values. Let the scalar used to normalize each image, i, be α_i. Corresponding output images are then also scaled by α_i to maintain a proper relationship between magnetic field amplitude and current density, given finite computational precision. The network performs best when a separate, global scaling factor β ($\sim 10^8$) is applied to all output data before training; and the inverse of this factor $1/\beta$ is applied afterwards to the MAGIC-UNet prediction. Given that β affects average values by many orders of magnitude, this operation helps ensure that the mean pixel value of input data is not so small or large as to limit the flexibility of the network's activation function. Between every epoch of the training process, additive Gaussian noise drawn from a random seed is applied to the input data. This noise is found to be a reasonable match to the noise behavior of our experimental data.

To evaluate the performance of the trained MAGIC-UNet network on out-of-distribution data, we create sets of input-output pairs, generating each set with different methods. In one set, we simulate random, straight-current densities directly in COMSOL® using an alternative algorithm and with separate parameters to the synthetic Class II data described above. We then obtain predicted magnetic field amplitudes from these current densities using COMSOL®. These simulated magnetic field images not only serve as effective, out-of-distribution outputs, but also as a method to validate the Forward Fourier Method calculation for creating training data. Separately, we generate custom current densities in COMSOL® that mimic current flow through devices with branched current and nonuniform wire width to test MAGIC-UNet performance on data far outside of the distribution of training data.

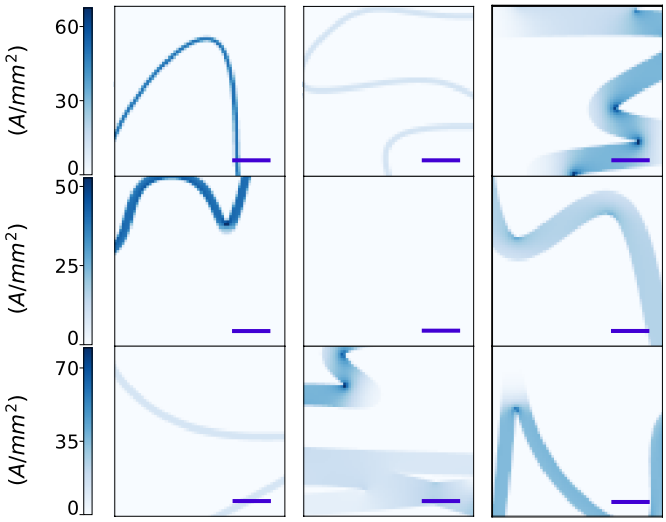

FIG. S3. Examples of current density magnitude from Class I data. Current flow around bends and overlapping wires is simulated using COMSOL Multiphysics® to realistically simulate current behavior around corners. Scale bars are 400 μm for all planar images shown.

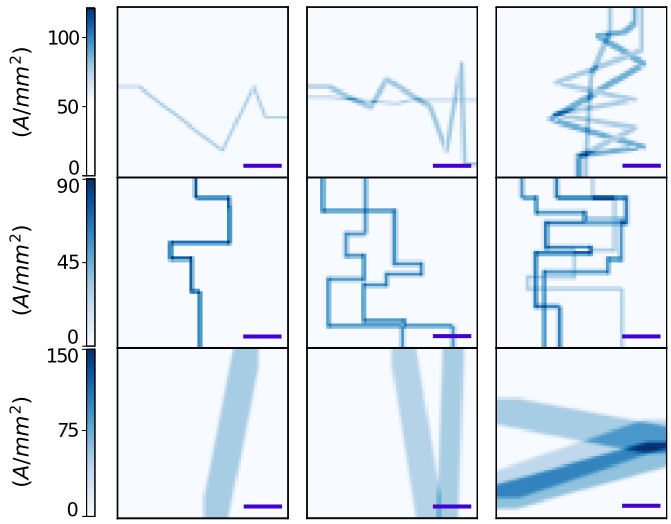

FIG. S4. Example current density magnitudes from Class II data. This data includes straight wire segments at different angles, and can include 1, 2, or 3 closely spaced current planes. Scale bars are 400 μm for all planar images shown.

Parameter	Class I	Class II
Class of wire	Curves formed by interpolating 2-21 randomly generated points (with cropping)	Thin right angle segments, thin arbitrary angle segments, thick straight wires
Number of independent wires	1	1, 2, or 3
Field of view	2 mm	2 mm
Standoff distance	50 $\pm$ 10 μm (500 $\pm$ 50 μm for large standoff)	50 $\pm$ 10 μm (500 $\pm$ 50 μm for large standoff)
Thickness of current layer	14 µm	14 µm
Width of wires (% of data set)	16-160 μm (50%) 160-320 μm (50%)	9-30 µm (67%) 97-156 µm (33%)
Approximate Current Range	15.5 - 311 mA	1.95 - 170 mA

TABLE S1. Parameters for synthetic current density data generation, used for MAGIC-UNet training and validation.