(Form 5)

Forefront Physics and Mathematics Program to Drive Transformation (FoPM)

Report for the International Research Experience

Name	Wang Huidong
Affiliation	Department of Chemistry, Graduate School of Science
Host Institution	Anhui Jianzhu University, School of Mechanical & Electrical Engineering
Period	From 07/01/2024 to 08/31/2024

During my international research experience at Anhui Jianzhu University, I worked on imaging through scattering media using deep learning. The motivation is easy to state but hard to solve: when light passes through a turbid or rough medium—fog, frosted glass, or biological tissue—it is scattered into a disordered "speckle" pattern that the eye cannot interpret. Recovering, or at least recognizing, an object hidden behind such a medium is a long-standing problem in optics with real practical value, from imaging in poor visibility to seeing through tissue. My goal during the stay was to understand how computational methods, and machine learning in particular, can turn these scrambled patterns back into useful information.

I began by building and characterizing the optical system used to generate the data. A light source illuminated an object in the object plane—I used handwritten-digit images from the MNIST dataset as stand-in objects so that every measurement carried a known label—after which the light passed through a scattering diffuser and was recorded as a speckle pattern on a camera. Working with a 633 nm source and fixed lens geometry, I collected several thousand speckle images through one scattering medium and a smaller set through a second, different medium. Setting this up taught me a great deal about practical Fourier optics, optical alignment, and the trade-offs between signal, noise, and acquisition time.

A recurring difficulty in this field is that deep-learning models need far more labelled data than an optical bench can comfortably produce. My first task was to address this without collecting more measurements. Using the Gerchberg–Saxton phase-retrieval algorithm, I recovered the hidden object from the autocorrelation of a single speckle pattern, and then estimated the system's point-spread function (PSF) by deconvolution. Because each recovered PSF differs from the others essentially only by noise—I verified their similarity using the structural similarity index, which stayed within a narrow, high range—I could convolve recovered objects with these PSFs to synthesise new, physically plausible speckle images. In effect, I expanded a small experimental dataset into one large enough to train a network, without injecting any external information. Implementing the iterative phase retrieval, including the hybrid input–output and error-reduction steps for noise suppression, gave me a concrete understanding of how phase information—which the camera cannot directly record—can be reconstructed.

With a usable dataset in hand, I compared two very different learning approaches. The first was a support vector machine (SVM), which I used to classify speckle patterns directly into digit labels. Deriving the formulation from the maximum-margin principle through the KKT conditions and the Lagrangian dual deepened my understanding of classical machine learning. The SVM proved remarkably data-efficient and accurate when the training and test speckles came from the same scattering medium, reaching strong classification accuracy with only a modest number of training

images. Its key limitation, which became the central lesson of the project, was that it could not generalise: a model trained on one medium performed very poorly on speckles from a different medium, because it had learned the labels rather than any transferable physical relationship.

To overcome this, I turned to convolutional neural networks (CNNs). A standard time-domain CNN struggled with our limited dataset, partly because conventional max-pooling discards a large fraction of the information at each downsampling step. I therefore implemented a frequency-domain CNN based on spectral pooling, which performs downsampling in the Fourier domain and so preserves more fine structure as information flows between layers. This change was decisive: the spectral-pooling network trained several times faster, tolerated the small dataset, and—most importantly—a model trained only on the first scattering medium was able to classify speckles formed by a second, previously unseen medium with encouraging accuracy. This generalisation across different scattering systems was the outcome I was most pleased with, since it points toward a single model that could handle a range of optical conditions rather than being rebuilt for each one. By modifying the output layer, the network could also move beyond labels toward reconstructing the shape of the hidden object directly.

The experience taught me as much about how to do research as about the science itself. I learned to move fluidly between physical modelling and data-driven methods and to treat them as complementary: the phase-retrieval and PSF model supplied the physical prior that made the learning problem tractable, while the network supplied the generalisation that the classical pipeline lacked. On the technical side, I became comfortable implementing algorithms in both MATLAB and Python, building and debugging a TensorFlow training pipeline, and reasoning carefully about overfitting, leakage between media, and evaluation metrics such as precision, recall, and F1-score. Just as valuable were the softer lessons—framing a tractable question, iterating when an approach failed, and presenting results clearly.

Looking back, this project shaped the direction of my later research, in which I again combine optics, imaging, and machine learning. The international setting added another dimension: working in an unfamiliar academic environment, communicating across a language barrier, and adapting to a different research culture were challenges in themselves, and overcoming them gave me confidence that I could contribute in a new setting. I am grateful for the guidance of my host and colleagues, and I left with both a concrete set of skills and a clearer sense of the kind of researcher I want to become.



Figure 1. The optical setup used to acquire the speckle data. The components—light source, lenses, the scattering medium, and the camera—are mounted on a vibration-isolated optical table, with the beam path arranged so that light from the source illuminates the object plane, passes through the scattering diffuser, and is recorded as a speckle pattern.