

機械学習を活用したイメージフリーシングルピクセル物体認識

Image-free single-pixel object recognition using machine learning

研究代表者 金沢大学理工研究域 助教 遠藤優

Yutaka Endo

Single-pixel imaging (SPI) is a computational imaging technique to obtain images using a single-pixel detector. SPI is expected to be used in terahertz imaging, bio-imaging, and 3D sensing, where it is difficult to use 2D image sensors. Recently, *image-free* SPI has been studied, which performs high-level vision tasks (e.g., object classification, detection, recognition, etc.) from measurements obtained by an SPI system without image reconstruction. Image-free SPI is expected to realize a measurement system for fast object recognition. In this study, we developed a method to jointly optimize SPI patterns and a classifier using machine learning to realize a fast and accurate Image-free SPI system. SPI patterns are important design parameters of the SPI system, and optimized SPI patterns and classifier enable fast and accurate object recognition. We tested our method on two datasets: synthetic data obtained by simulation and real data obtained by an SPI optical setup. The experimental results show that our jointly learned patterns and classifier outperforms conventional method using Hadamard and random patterns.

要旨

シングルピクセルイメージング (SPI: single-pixel imaging) は単一画素検出器を用いて画像を取得する手法である。SPI は多数の画素を持つイメージセンサの利用が困難な場面での活用が期待されており、テラヘルツイメージ、バイオイメージング、3D センシングなどに応用されている。また近年、SPI システムを用いて取得したデータから、物体の画像を再構成することなく、物体認識を行う「イメージフリーSPI」が注目されている。イメージフリーSPI を用いれば高速に物体認識ができる計測システムを実現できると期待されている。本研究では、イメージフリーSPI に基づく高速物体認識の実現のために、機械学習を用いた変調パターンの最適設計に取り組んだ。変調パターンは SPI における計測性能を決める重要な設計パラメータであり、これを最適化することで、高速・高精度な物体認識が可能となる。学習したパターンを使って計測を行うことで、従来用いられているアダマールパターンやランダムパターンと比べて、少ない計測回数で高精度な分類が可能であることを示した。

1. まえがき

画像の取得とその解析は、マシンビジョンやライフサイエンスなど幅広い領域で用いられる重要な要素技術である。CCD や CMOS プロセス技術の発達により、身の回りのカメラに数千万画素のイメージセン

サが搭載され、手軽に大容量な画像取得が可能になっている。また近年の機械学習関連技術の飛躍的発展により、画像解析技術も大きな発展を遂げ、監視カメラ、製品検査、医療診断、車載システムなどへ応用されている。一方で、波長帯、感度、応答性、データ量などの制約により、高密度な画素アレイを持つイメージセンサの使用が困難な応用も数多く存在する。例えば、テラヘルツ波や近赤外線などの波長帯は製品検査や試料分析などに用いられるが、そのような波長帯では画素アレイが高コストであることが多い。また微弱な光を高感度に捉えることは顕微鏡技術などで重要であるが、そのような用途では光電子増倍管や SPAD (single photon avalanche diode) などの高感度なフォトディテクタ用いられる。また高速な応答性は高速現象の観察や LiDAR (light detection and ranging) などで求められるが、一般に CCD や CMOS よりもフォトディテクタの方が高速に応答する。そのため、そのような制約の厳しい環境下で画像を取得・解析できる技術が求められている。

シングルピクセルイメージング (SPI: single-pixel imaging) は単一画素検出器を用いて画像を取得する手法である^(1,2,3)。SPI は上記のような多数の画素を持つイメージセンサの利用が困難な場面での活用が期待されており、テラヘルツイメージング^(4,5,6)、バイオイメージング^(7,8,9)、3D センシング^(10,11,12) などに応用されている。また近年、SPI システムを用いて取得したデータから、物体画像を再構成することな

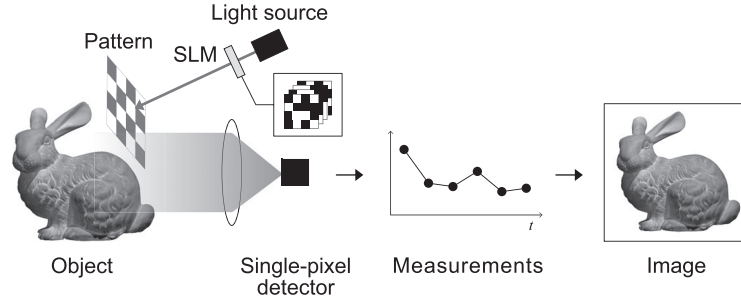


図1 シングルピクセルイメージング

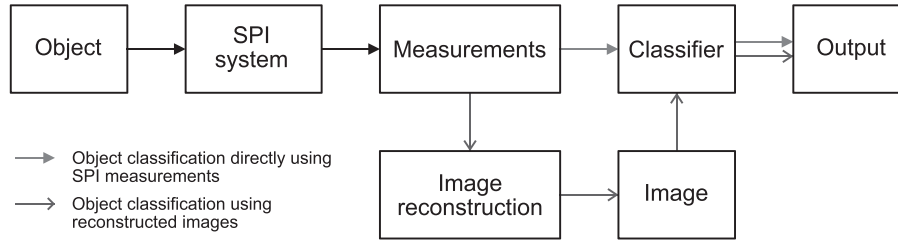


図2 SPI システムを用いた物体認識

く物体認識を行う「イメージフリーSPI」の研究が注目されている^(9, 13, 14, 15, 16)。SPI で高精細な画像を再構成するには多くの計測データ (長い計測時間) が必要である。必要な出力が画像そのものではなく認識結果である場合、画像再構成を行わなければ、少ない計測データ数で高速に結果を得ることができる。そのため、イメージフリーSPI を用いれば高速に物体認識を行える計測システムを実現できると期待されている。このようなアプローチに基づき、画像分類^(13, 14, 15)、細胞分類⁽⁹⁾、物体位置検出⁽¹⁶⁾などが行われている。

本研究では、イメージフリーSPI に基づく高速物体認識の実現のために、機械学習を用いた変調パターンの最適設計に取り組んだ。変調パターンは SPI における計測性能を決める重要な設計パラメータである。最適な変調パターンを用いることで、より少ない計測回数で (高速に) 高精度な物体認識が可能となる。特定の認識タスクに有用なパターンの設計は人手で行うことが難しいが、機械学習を用いることで、データ駆動でパターンを設計することが可能となる。提案手法をシミュレーション及び実際のシステムを用いた光学実験で検証し、従来よりも少ない計測回数で高精度な物体認識が可能であることを実証した。

2. シングルピクセルイメージング

図1にSPIによる計測と画像再構成の概略を示す。光源からの光をSLM (spatial light modulator) を用いて変調し、照明パターンを形成する。パターンを対象物体に照射し、その反射光または透過光をフォトディテクタに集光し、その強度を計測する。異なるパターンを用いてこの計測を繰り返す、得られた計測データから対象を再構成する。 N 画素からなる対象物体を $\mathbf{x} \in \mathbb{R}^N$ 、 i 回目に照射したパターンを $\mathbf{w}_i \in \mathbb{R}^N$ とすると i 回目に計測された光強度 y_i は

$$y_i = \mathbf{w}_i^T \mathbf{x}$$

となる。異なる M 枚のパターンを用いて計測した時の計測データを $\mathbf{y} = [y_1 \ \cdots \ y_M]^T \in \mathbb{R}^M$ とすると

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (1)$$

となる。ここで $\mathbf{W} \in \mathbb{R}^{M \times N}$ は i 行ベクトルが $\mathbf{w}_i^T$ となる行列である。この線形方程式系で表される逆問題を解くことで、対象 $\mathbf{x}$ を再構成できる。計測回数が $M < N$ のとき逆問題は劣決定となり、一意の解を得られない。そのため、基本的には再構成したい画像の画素数 N と同じだけの計測回数 M が必要となる。よって、画素数が多い物体を再構成したい場合には、多数の計測回数が必要となる。

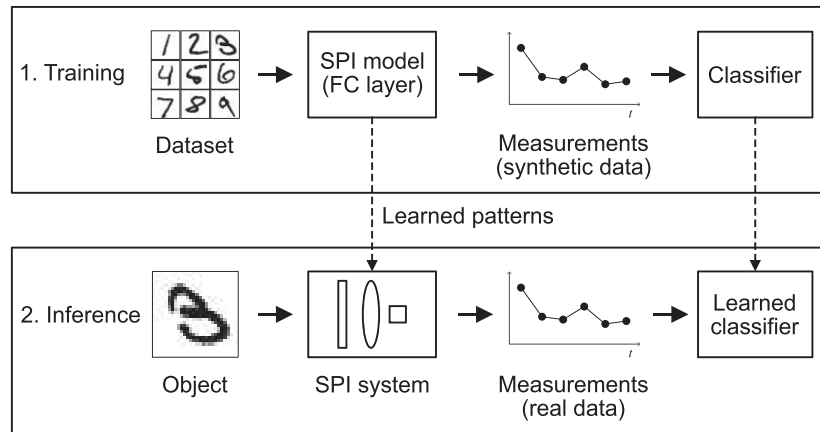


図3 SPI パターンと分類器の同時最適化

3. SPI システムを用いた物体認識

3.1 イメージフリーSPI 物体認識

近年, SPI システムを用いて物体認識を行う研究が盛んに行われている^(9, 13, 14, 15, 16). 図2にSPI システムを用いた物体認識の概略を示す. 画像を用いた物体認識では, 画像データセットを用いて分類器を学習し, 学習済み分類器に分類したい画像を入力して推論を行う. SPI システムを用いて画像認識を行う場合も同様に, 画像データセットを用いて分類器を学習し, SPI システムで取得した計測データ $\mathbf{y}$ から再構成した画像 $\mathbf{x}$ を学習済み分類器へ入力して分類を行うことができる. ただし, この方式ではSPI の計測データから画像再構成をする必要があり, その分の計算時間が必要となる. また分類に高精細な画像が必要になる場合, 画素数 N に応じた多数の計測データ数 M が必要となる. 一方, 計測データ $\mathbf{y}$ を入力データとして用いて分類器を学習しても, 画像認識が可能であることが示されている^(17, 18). この方式では画像再構成が不要となり, 計算時間を削減できる. また K クラスの分類を行う場合, クラス数 K に応じた数の計測データ数 M が必要となる. よって, $K < N$ であれば, 画像再構成が必要なときよりも少ない計測回数で分類できる.

今回, 手書き数字画像のデータセットである MNIST と衣料品画像のデータセットである Fashion-MNIST を用いた分類問題をターゲットに SPI での物体認識に取り組む. MNIST は 0~9 の手書き数字画像を集めたデータセットであり, トレーニングデータ 50,000 枚, テストデータ 10,000 枚で構成される. Fashion-MNIST は T シャツや靴などの 10 種類の衣料品画像を集めたデータセットであり, トレーニングデータ

60,000 枚, テストデータ 10,000 枚で構成される.

3.2 SPI パターンと分類器の同時最適化

SPI システムで用いられるパターンは計測性能を決める重要な設計パラメータである. SPI で主に用いられているパターンは, アダマールパターンやランダムパターンである. このパターンを機械学習により最適化することで, 少ない計測データ数で高精細な画像を再構成できることが報告されている^(12, 19). 本研究では, SPI システムを用いた物体認識において使用するパターンと分類器を機械学習によって同時に最適化する手法を提案する.

図3にパターンと分類器の同時最適化の概略を示す. 式(1)で表される SPI の計測モデルは, 対象物体画像を入力, 変調パターンを重み, 計測データを出力としたニューラルネットワークの全結合層とみなせる. よって, SPI システムの計測データを用いて分類を行うことは, 全結合層 (SPI モデル) と分類器からなる機械学習モデルに対象物体画像を入力していると見なすことができる. このモデルを計測対象 $\mathbf{x}$ に相当するデータセットを用いて訓練することで, パターンと分類器を同時に最適化できる. 実際の SPI システムを用いて推論を行う際には, 学習されたパターンを用いて得られた計測データを, 学習済みの分類器に入力することで分類結果を得ることができる.

3.3 DMD に適したパターンの設計

本研究では SLM として DMD (digital micromirror device) を使用する. DMD は多数の微小ミラーがアレイ状に配置されたデバイスである. 各ミラーは角度を ON/OFF の 2 値に制御することができる. DMD は 2

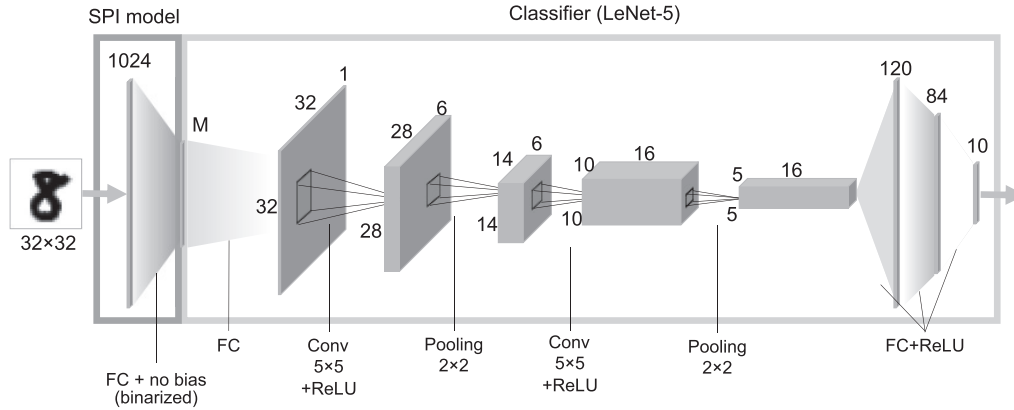


図4 SPI パターンと分類器の同時最適化に用いたモデル

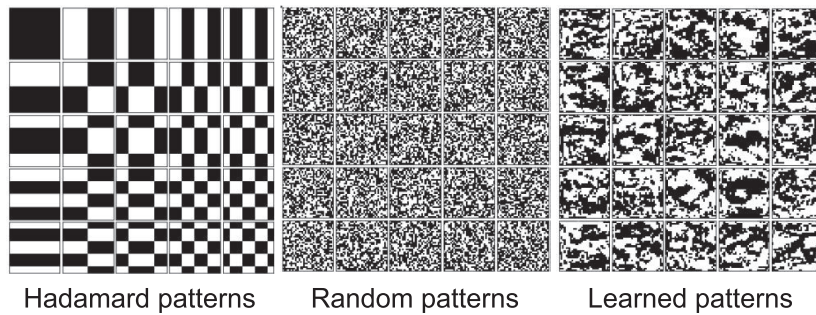


図5 SPI パターンと分類器の同時最適化

値（バイナリ）の振幅変調 SLM として用いられている。DMD を用いた計測モデルでは、 $\mathbf{W}$ の要素として $\{0, 1\}$ の値しか取れない。そこでモデルの学習時に、SPI モデルの重みの2 値化を行った。DMD を用いた計測モデルと同じ $\{0, 1\}$ の2 値化処理を行った場合、重みの値が 0 になることで学習が停滞してしまうことがわかった。そのため、Binarized Neural Network で提案されている手法⁽²⁰⁾と同様に、2 値化処理を $\{-1, +1\}$ で行った。 $\{-1, +1\}$ で2 値化されたパターン $\mathbf{W}$ を使用するときは、 $\mathbf{W}$ の要素のうち+1 以外を 0 としたパターン $\mathbf{W}_+$ 、-1 以外を 0 としたパターン $\mathbf{W}_-$ に分解して使用する。それぞれのパターンで得られた計測データを $\mathbf{y}_+$ 、 $\mathbf{y}_-$ とすると、計測データ $\mathbf{y}$ はそれらの差分 $\mathbf{y}_+ - \mathbf{y}_-$ によって得ることができる。

3.4 モデルの学習

図4 に本研究で使用した機械学習モデルを示す。SPI モデルとしてバイアスなしの全結合層を使用し、分類器として LeNet-5 をベースにモデルを構築した。LeNet-5 を用いた理由は、このモデルが比較的単純な構造であるため、パターンによる分類精度の違いが分かりやすいと考えたためである。入力データは後

述する光学実験の都合上、白黒反転し、 32×32 画素にリサイズした。学習の際の損失関数として交差エントロピー誤差を使用した。また最適化は Adam を用いて行った。図5 に学習されたパターンを示す。従来使用されているアダマールパターン、ランダムパターンとは異なるパターンが学習されていることがわかる。

4. シミュレーションによる分類精度の検証

シミュレーションにより画像データセットから SPI システムの計測データを合成し、分類精度を検証した。比較対象として、SPI で広く用いられているアダマールパターンとランダムパターンを用いて分類器の学習を行った。また計測回数 M による精度の変化を確認するため、 $M = 16$ の場合と $M = 64$ の場合で学習されたパターンを用いて精度を検証した。

表1 にシミュレーションで生成した合成データにおける分類精度を示す。MNIST, Fashion-MNIST の両方において、学習パターンを使用したとき最も認識精度が高くなった。また計測回数 M が小さいとき、アダマールパターン、ランダムパターンは精度が大きく減少しているのに対して、学習パターンでは精度

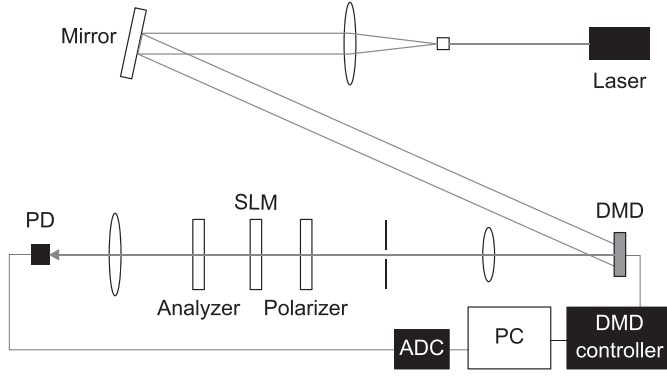


図6 開発したSPI システム

の減少幅が小さくなった。以上より、学習パターンの有効性をシミュレーションによって示すことができた。

表1 合成データに対する分類精度

Patterns	M	Dataset	
		MNIST	Fashion MNIST
Hadamard	16	85.6%	79.4%
	64	96.7%	86.4%
Random	16	83.5%	78.1%
	64	95.5%	83.2%
Learned	16	96.2%	85.3%
	64	97.9%	87.0%

5. 光学実験による分類精度の検証

5.1 SPI システムの実装

実際の SPI システムで取得された実データを用いて検証を行うため、図6のようなSPI 光学系を構築した。光源にはHeNe レーザ、SLMにはDMD (V-7001VIS, VIALUX)、単一受光素子としてSi 増幅フォトディテクタ (Thorlabs, PDA100A2) を使用した。計測対象には透過型液晶ディスプレイ (LC2012, Holoeve) に表示した画像データを使用した。光源からの光のビーム径を拡大し、DMD に照射する。DMD にはパターンを表示し、計測対象へパターンを投影する。計測対象からの透過光の光強度をフォトディテクタで計測する。この SPI システムとアダマールパターンを用いて計測したデータから再構成した画像を図7に示す。

5.2 実データを用いた分類精度の検証

図6のSPI システムで取得した実データを用いて分類精度の検証を行った。シミュレーションでの精度検証と同様に、学習パターン、アダマールパターン、ランダムパターンを用いて計測を行った。計測対象



図7 SPI システムで取得した再構成画像

として、MNIST と Fashion-MNIST データセットのテストデータからそれぞれランダムに 1,000 枚の画像を選出し、透過型液晶ディスプレイに表示した。

表2に実データにおける分類精度を示す。シミュレーションで生成した合成データにおける結果と同様に、MNIST, Fashion-MNIST の両方において、学習パターンの分類精度が最も高くなった。また計測回数 M が小さいとき、アダマールパターン、ランダムパターンは精度が大きく減少しているのに対して、学習パターンでは精度の減少が小さくなった。一方、すべてのパターンにおいて、実データを用いた場合の方が合成データを用いた場合と比べて分類精度が低下した。これは図7からも示唆される通り、実計測時のノイズ、背景光、位置ずれなどの影響で、SPI システムで取得される実データは合成データと性質が異なっているためだと考えられる。

表2 実データに対する分類精度

Patterns	M	Dataset	
		MNIST	Fashion MNIST
Hadamard	16	45.0%	56.3%
	64	83.8%	78.8%
Random	16	50.6%	56.4%
	64	67.0%	72.0%
Learned	16	90.0%	79.8%
	64	92.0%	81.4%

5.3 データミスマッチへの対応

画像データセットを用いた学習では、式(1)に従って計測データ $\mathbf{y}$ が合成される。実際のSPI システムで取得される実データは、ノイズ、背景光、位置ずれなどの影響を受けるため、合成データと実データの間にミスマッチが生じ、分類精度が低下する。このデータミスマッチを防ぐために、データ拡張とファイ

チューニングを試みた。データ拡張ではデータセットの画像に対し、回転、平行移動、拡大、縮小の操作を加え、さらにガウスノイズを付加した。ファインチューニングでは、(1)合成データを用いてパターンと分類器を学習し、(2)学習したパターンを用いて実データを計測し、(3)実データを用いて分類器のみを再学習する。

データミスマッチへの対応方法として、(1)データ拡張、(2)ファインチューニング、(3)データ拡張＋ファインチューニングの3通りの手法を検証した。MNIST と Fashion-MNIST データセットのトレーニングデータからそれぞれランダムに10,000枚の画像を選出し、SPI 光学系で計測した実データをファインチューニングに用いた。表3にデータ拡張とファインチューニングを行ったときの実データの分類精度を示す。データ拡張、ファインチューニングを行うことで、どちらも実データの認識精度を向上させられることがわかった。このことからデータ拡張とファインチューニングがデータミスマッチの解消に有効であることを示した。

表3 データ拡張とファインチューニングを用いたときの実データに対する分類精度

	M	Dataset	
		MNIST	Fashion MNIST
Augmentation	16	92.2%	85.2%
	64	92.0%	80.9%
Fine tuning	16	96.3%	89.1%
	64	97.4%	89.4%
Augmentation + Fine tuning	16	96.5%	90.1%
	64	98.4%	91.8%

6. まとめ

イメージフリーSPIにおける変調パターンを機械学習によって最適設計する方法を提案した。提案手法では、SPIモデルと分類器からなるニューラルネットワーク訓練することで、パターンと分類器を同時に最適化する。提案手法を手書き数字認識(MNIST)と衣料品画像分類(Fashion-MNIST)に適用し、その有効性を検証した。シミュレーションによって得られた合成データと実際のSPIシステムで得られた実データを用いて分類精度を検証したところ、従来用いられているアダマールパターン、ランダムパターンと比べて高精度な分類を達成でき、その有効性を実証できた。また合成データと実データとの間にミ

スマッチがあることが実験により確認された。このデータミスマッチを解消するため、データ拡張とファインチューニングを試み、その有効性を示した。今後は、さらにクラスの多い問題や実用的なデータセットに対して本手法の有効性を検証していく。

発表論文

- [1] 中島 凱, 遠藤 優, 単一画素カメラにおける空間変調パターン
の物体認識に対する最適化, 電子情報通信学会 2020 年総大会, ISS-SP-023, 2020 年3月.
- [2] 中島 凱, 遠藤 優, 物体認識タスクに最適化した単一画素カメラの性能評価, Optics & Photonics Japan 2020, 17aB9, 2020 年11月.

参考文献

- (1) M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, Ting Sun, K. F. Kelly, and R. G. Baraniuk, "Single-Pixel Imaging via Compressive Sampling," IEEE Signal Process. Mag. 25, 83-91 (2008).
- (2) M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," Nature Photon 13, 13-20 (2019).
- (3) G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: a review," Opt. Express 28, 28190-28208 (2020).
- (4) C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith, and W. J. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," Nature Photon 8, 605-609 (2014).
- (5) R. I. Stantchev, B. Sun, S. M. Hornett, P. A. Hobson, G. M. Gibson, M. J. Padgett, and E. Hendry, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," Sci. Adv. 2, e1600190 (2016).
- (6) R. I. Stantchev, X. Yu, T. Blu, and E. Pickwell-MacPherson, "Real-time terahertz imaging with a single-pixel detector," Nat Commun 11, 2535 (2020).
- (7) V. Studer, J. Bobin, M. Chahid, H. S. Mousavi, E. Candes, and M. Dahan, "Compressive fluorescence microscopy for biological and hyperspectral imaging," Proceedings of the National Academy of Sciences 109, E1679-E1687 (2012).
- (8) K. Guo, S. Jiang, and G. Zheng, "Multilayer fluorescence imaging on a single-pixel detector," Biomed. Opt. Express 7, 2425 (2016).
- (9) S. Ota, R. Horisaki, Y. Kawamura, M. Ugawa, I. Sato, K. Hashimoto, R. Kamesawa, K. Setoyama, S. Yamaguchi, K. Fujiu, K. Waki, and H. Noji, "Ghost cytometry," Science 360, 1246-1251 (2018).
- (10) B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, "3D Computational Imaging

- with Single-Pixel Detectors," *Science* 340, 844-847 (2013).
- (1 1) M.-J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett, "Single-pixel three-dimensional imaging with time-based depth resolution," *Nat Commun* 7, 12010 (2016).
 - (1 2) N. Radwell, S. D. Johnson, M. P. Edgar, C. F. Higham, R. Murray-Smith, and M. J. Padgett, "Deep learning optimized single-pixel LiDAR," *Appl. Phys. Lett.* 115, 231101 (2019).
 - (1 3) S. Lohit, K. Kulkarni, P. Turaga, J. Wang, and A. C. Sankaranarayanan, "Reconstruction-free inference on compressive measurements," in 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (IEEE, 2015), pp. 16-24.
 - (1 4) S. Jiao, J. Feng, Y. Gao, T. Lei, Z. Xie, and X. Yuan, "Optical machine learning with incoherent light and a single-pixel detector," *Opt. Lett.* 44, 5186 (2019).
 - (1 5) P. Latorre-Carmona, V. J. Traver, J. S. Sánchez, and E. Tajahuerce, "Online reconstruction-free single-pixel image classification," *Image and Vision Computing* 86, 28-37 (2019).
 - (1 6) Z. Zhang, J. Ye, Q. Deng, and J. Zhong, "Image-free real-time detection and tracking of fast moving object using a single-pixel detector," *Opt. Express* 27, 35394 (2019).
 - (1 7) M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal Processing With Compressive Measurements," *IEEE J. Sel. Top. Signal Process.* 4, 445-460 (2010).
 - (1 8) R. Calderbank and S. Jafarpour, "Finding needles in compressed haystacks," in *Compressed Sensing*, Y. C. Eldar and G. Kutyniok, eds. (Cambridge University Press, 2012), pp. 439-484.
 - (1 9) C. F. Higham, R. Murray-Smith, M. J. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," *Sci Rep* 8, 2369 (2018).
 - (2 0) M. Courbariaux, I. Hubara, D. Soudry, R. El-Yaniv, and Y. Bengio, "Binarized Neural Networks: Training Deep Neural Networks with Weights and Activations Constrained to +1 or -1," (2016).