

---

論文紹介 23/07/17

# Ultrafast single-channel machine vision based on neuro-inspired photonic computing

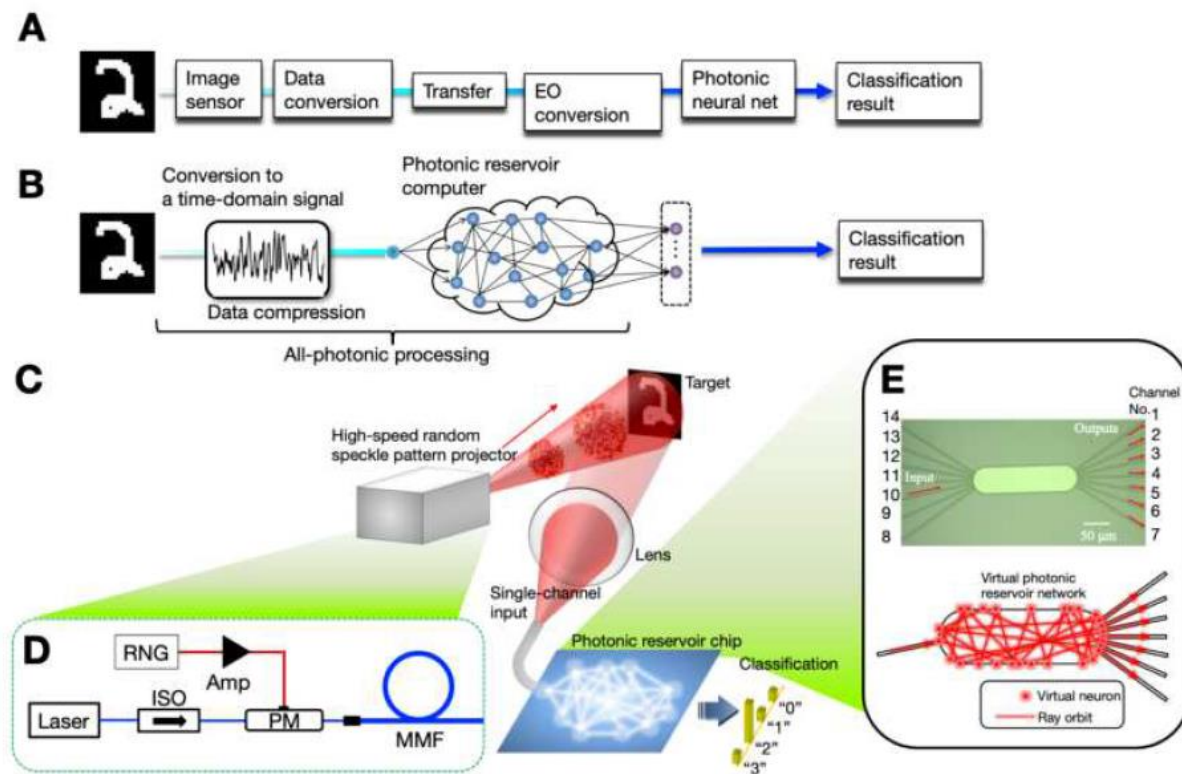
T. Yamaguchi, K. Arai, T. Niiyama, A. Uchida, and S. Sunada

明治大学 M1 大塚 雄太

# 概要

高速なマシンのビジョンは科学と技術の両分野でその重要性が高まっている。Neuro-inspired photonic computingは、超低遅延でマシンビジョンを高速化させる有望なアプローチである。しかし、イメージセンサーのフレームレートによって処理速度が数十Hzに制限されてしまいます。本論文ではイメージセンサーを用いないマシンビジョンのフレームワークを提案する。このフレームワークは、ランダムな時間エンコーディング技術に基づき、単一の入力チャネルのみで実世界の視覚情報を光学的に処理する。このアプローチは、従来のアプローチを凌駕するGHzの速度で単一チャネルによる視覚情報の圧縮的取得を可能とし、時間領域でフットニックリザバコンピューティングを用いた直接的な光処理を可能とする。提案するアプローチが高速画像認識と異常検出が可能であることを実験的に実証し、さらに高速イメージングに利用できることを示す。提案アプローチは多目的であり、サブナノ秒現象の追跡、制御、補足をふくむ幅広い応用に拡張可能である。

# 基本動作原理



- ・図1Aは、従来のイメージセンサーに基づくフォトニックマシンビジョンシステムの概念図である。
- ・図1Bは、1チャンネルの画像取得とフォトニックRCプロセッサに基づき提案されたアプローチである。
- ・図1Cは、提案システムの概念図である。
- ・図1Dは、高速マシンランダムスペckルパターン投影機のセットアップである。RNGは乱数発生器、ISOは光アイソレータ、PMは位相変調器、MMFはマルチモードファイバである。

- ・図1Eは、入出力導波路チャンネルに結合されたスタジアム型マイクロキャビティに基づくオンチップフォトニックRCプロセッサである。マイクロキャビティは光学的な観点から内部カオス多重散乱によって仮想的なランダム光ネットワークを形成することができる。本研究では、導波路チャンネル(No.10)からリザーバーに信号を入力し、5つの出力チャンネル(No.2~6)からRC出力信号を抽出し、後処理に用いた。

# 基本動作原理(従来のフォトリックマシンビジョンシステム)

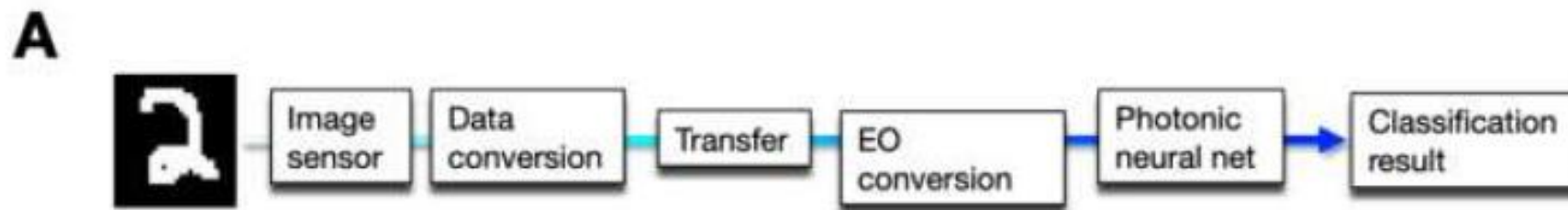
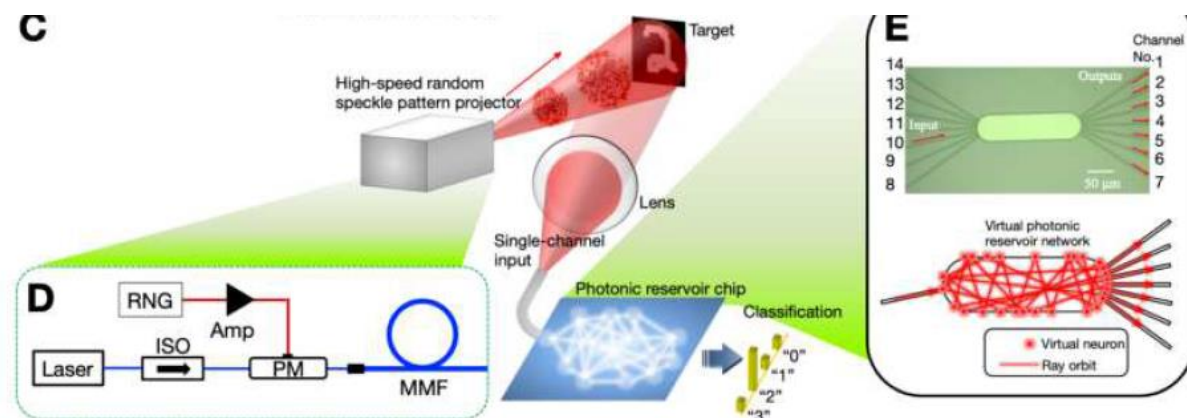


図1Aはターゲットの視覚情報を取得するために使用されるデジタルイメージセンサと、フォトリック処理ユニットを含む。システム全体の速度はこれによるデータ取得とデータ転送によって制限される(数十Hz程度)。また、多数の画素をもつイメージセンサを使用する場合は空間情報が電氣的領域に変換され、データ保存のために膨大な量のメモリが必要となる。

# 基本動作原理(今回提案するアプローチ)



提案されたシステムには、ターゲット物体の空間情報を時間的にエンコードするランダムパターンプロジェクタ(D)と、画像エンコードされた時間領域信号を処理するフォトニックRCプロセッサ(E)が含まれる。

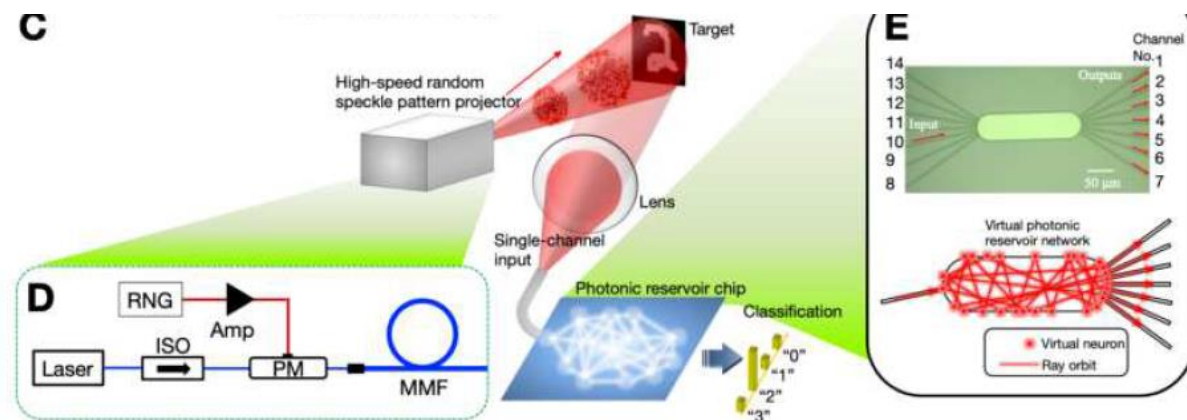
ランダムパターンプロジェクターは、ランダムなマスクパターンを生成し、それを対象物に照射する。光はターゲットから反射され、その像を $v(x, y)$ で表し、収束レンズによって集束され、直接フォトニックRCプロセッサに送られる。

$(x, y)$ は画像平面上の座標を表す。時間 $t$ における画像平面上のランダムマスクパターン $Mask(x, y, t)$ に対して、RCプロセッサへの入力光 $u(t)$ は、空間積分

$$\int Mask(x, y, t)v(x, y)dx dy$$

で与えられ、ターゲット画像の空間情報は時間領域の信号として符号化される。これにより $u(t)$ の特徴を高次元空間に個別に分布させることができ、その結果、簡単な後処理でより優れた画像認識を行うことができる。

# 基本動作原理(今回提案するアプローチ)



ここで、 $x_r(t)$ と $\phi(x_r(t)) \in R^M$ をリザーバの内部状態ベクトルとし、 $u(t)$ に応答する観測量とする。観測値 $\phi(x_r(t))$ は、取得時間 $T_N$ の間にサンプリング時間間隔 $\tau_S$ でサンプリングされる。

先行研究と同様に出カベクトル $y(n) \in R^{M_y}(n \in \{1, 2, \dots\})$ は、観測値 $\phi(x_r(t_j))$ 、読み出し重み $W(t_j) \in R^{M_y \times M}$ とバイアス $b \in R^{M_y}$ として、

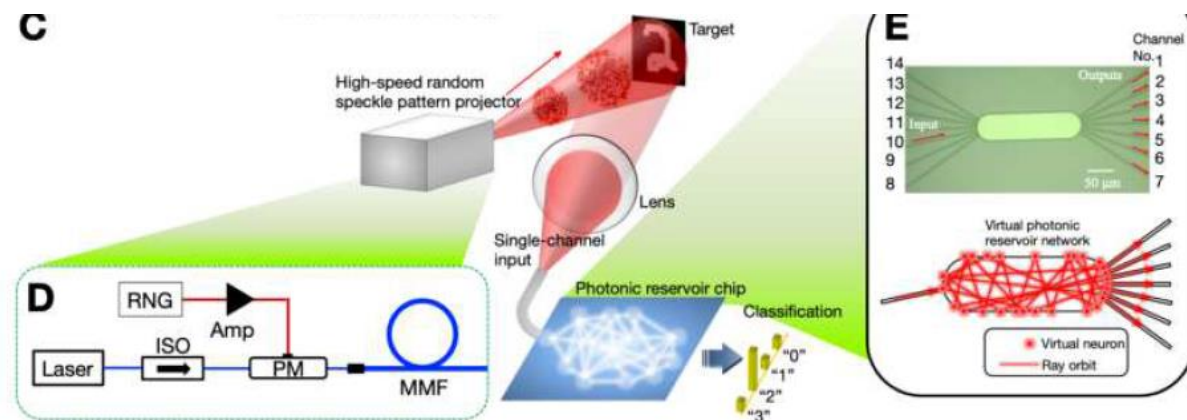
- ・回帰タスクでは、 $y(n) = \sum_{j=0}^{N-1} W(t_j) \phi(x_r(t_j)) + b$

- ・分類タスクでは、 $t_j = nT_N + j\tau_S$  ( $j \in \{0, 1, \dots, N\}$ )、 $N = T_N/\tau_S$ として、 $y(n) = f(\sum_j W(t_j) \phi(x_r(t_j)) + b)$ となる。 $f$ はソフトマックス関数である。

重み行列 $W(t)$ は、出力ベクトル $y(n)$ がターゲットベクトル $y_{tag}(n)$ に対応するように、訓練データセットを用いて訓練することが可能である。この方式では、時間間隔 $T_N$ で出力ベクトル $y(n)$ を得ることができる。



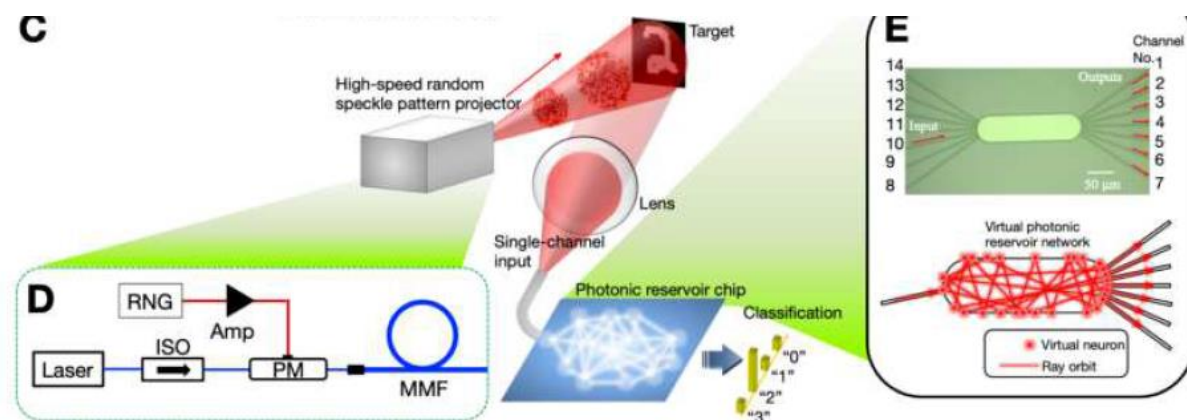
# 基本動作原理(今回提案するアプローチ)



標準的なリカレントニューラルネットワークと比較した場合のRCプロセッサの主な利点は、損失関数のグローバルミニマムにアクセスするための高速学習であり、その結果、学習のコストが低く抑えられることである。また、後処理は、アプリケーション固有の回路や、フィールドプログラマブルゲートアレイ(FPGA)を用いて低遅延で行うことができる。

本研究では、概念実証のため、RCプロセッサを用いた高速データ取得と前処理の能力評価に焦点をあてる。

# 基本動作原理(高速ランダムパターンプロジェクター)



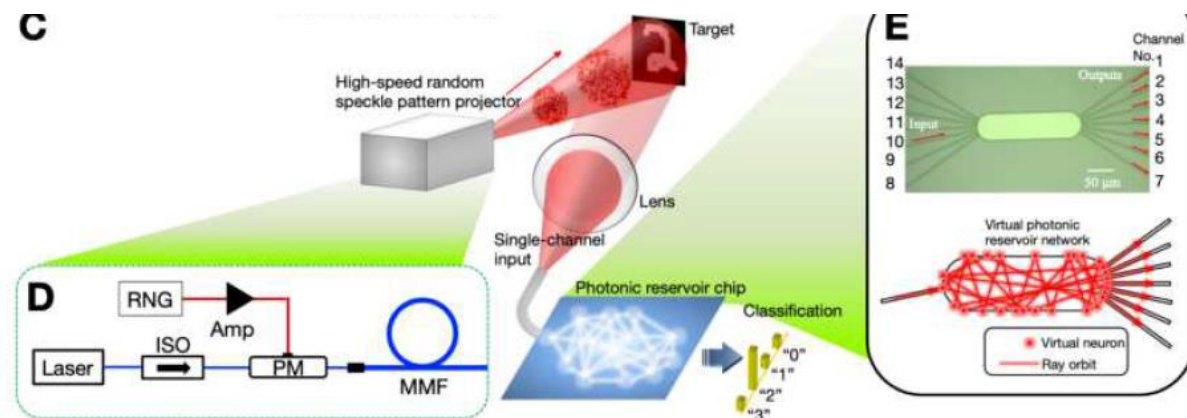
ランダムパターンプロジェクターは、先行研究の高速スペックルジェネレーターをベースとしており、レーザー光源、乱数発生器、位相変調器、マルチモードファイバから構成される。

コヒーレント光がMMFに入力されると、光は位相速度の異なる複数の伝搬モードに結合され、それらの干渉によって端面にスペックルパターンが形成される。スペックルは入射光の位相変化に敏感であるため、スペックルパターンは、入射光の位相を変調することによって変化させることができる。

今までの研究では、デジタルマイクロミラーデバイス(DMD)を含む空間光変調器(SLM)を用いて、最大でも22kHzのレートで光学ランダムスペックルパターンは生成されてきた。最近の研究では、機械的に回転するマスクパターンを用いて、最大2.4MHzの変調速度が実証された。一方、提案されているプロジェクターでは、位相変調率によって制御できる変調率は数十GHz以上になる。



# 基本動作原理(フォトニックリザーバーコンピューティングプロセッサ)

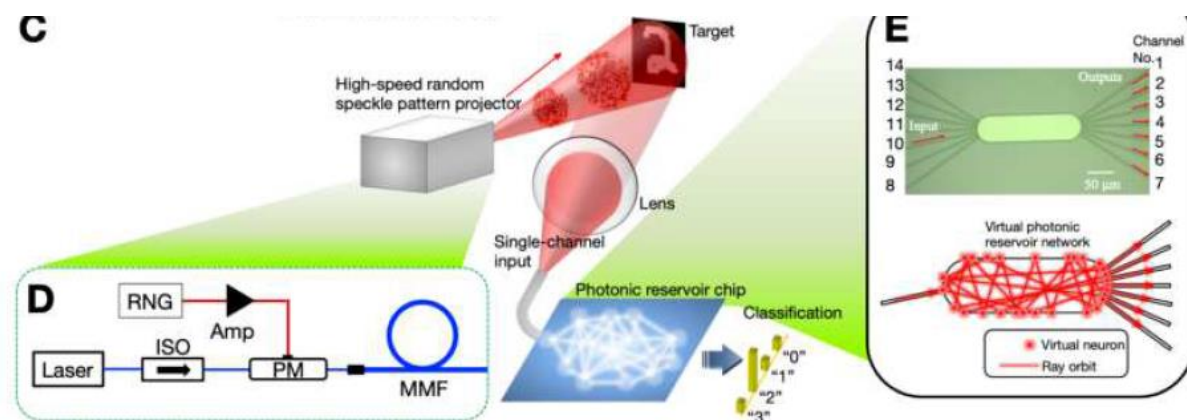


フォトニックRCプロセッサは、シリコンチップ上の14本のシングルモード導波路に結合されたスタジアム型マイクロキャビティ構造に基づいている。マイクロキャビティはリザーバーとして機能し、シングルモード導波路は入出力として使用される。

マイクロキャビティの特徴として、小さな設置面積で効率的に光を閉じ込めることができることと、マイクロキャビティの形状によって様々な波動パターンを形成できることである。

スタジアム型共振器は、Bunimovichスタジアムビリヤードにヒントを得た光線カオス共振器として知られる。カオス的性質による波の混合は、小さな設置面積で空間的に連続した光ランダムネットワークに対応する波動場を共振器内部に形成することができ、これは大規模なリザーバーとして機能する。

# 基本動作原理(フォトニックリザーバーコンピューティングプロセッサ)



非線形性は、強度検出と位相入力過程に自然に導入することができる。数値計算の結果、スタジアム型空洞ベースのRCは非カオティック空洞ベースのRCと比較して、非線形性を必要とするタスクに対して高い計算性能を持つことが明らかになった。

本研究で使用したスタジアム型RCチップの時間信号処理能力については、補論及び図S1~S3を参照。そこでは、フォトニックRCチップが、ベンチマークデータセットに対して、フォトニックRCシステムまたはフォトニックニューラルネットワーク回路を凌駕できることが示されている。

# 画像認識

提案システムの画像認識性能を評価した。実験では $28 \times 28$ ピクセルのMNIST手書き数字画像0~3を対象画像とし、DMDに表示した。ランダムスペckルパターンを生成し、25GS/sのレートでターゲットに投影した。反射光は光ファイバーを介してフォトニックRCプロセッサに入力された。RC出力は高速応答の光検出器で測定された。

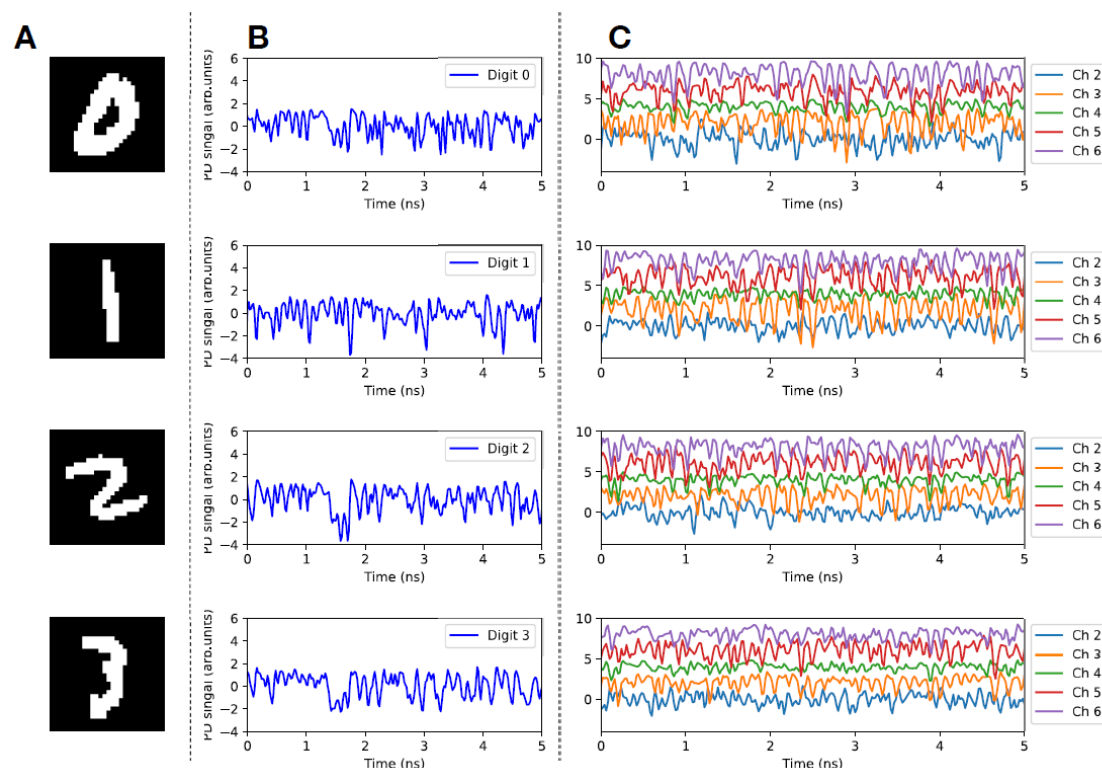


図2は、ターゲットの画像から反射された光の強度の時間変化と、それに対応するチャンネルNo.2~6からのRC出力を示している。反射光の波形はターゲットの画像に強く依存し、リザーバ出力の時空間応答は様々なものが生成される。

評価には、0から3までの数字画像を1000サンプル用い、各画像の取得時間 $T_N$ の間にRC出力を得た。予測出力 $y$ は、900の画像サンプルで学習され、100のサンプルでテストされた。

# 画像認識

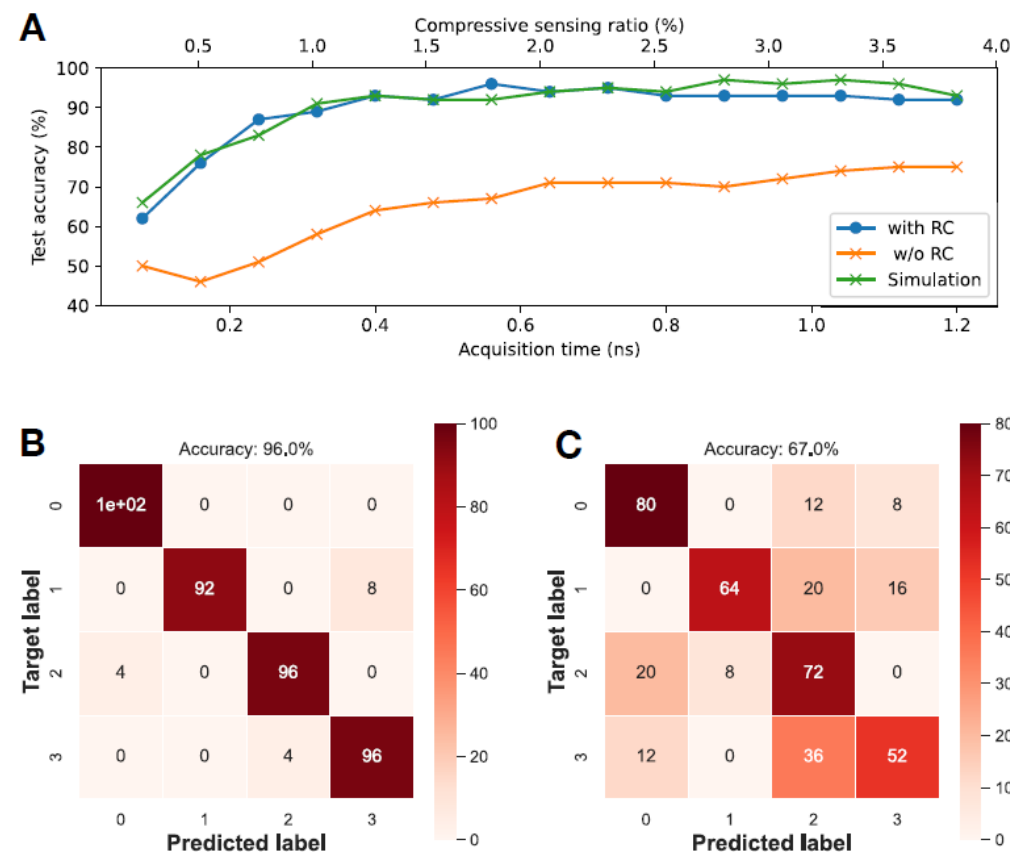


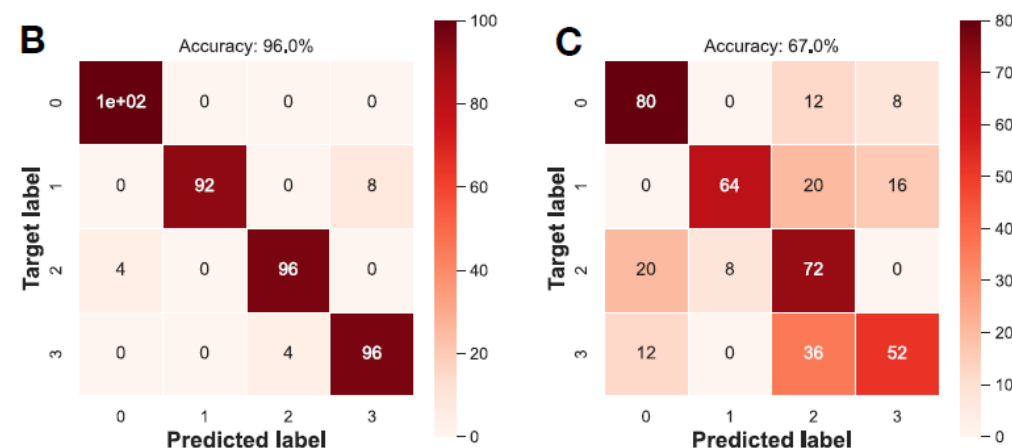
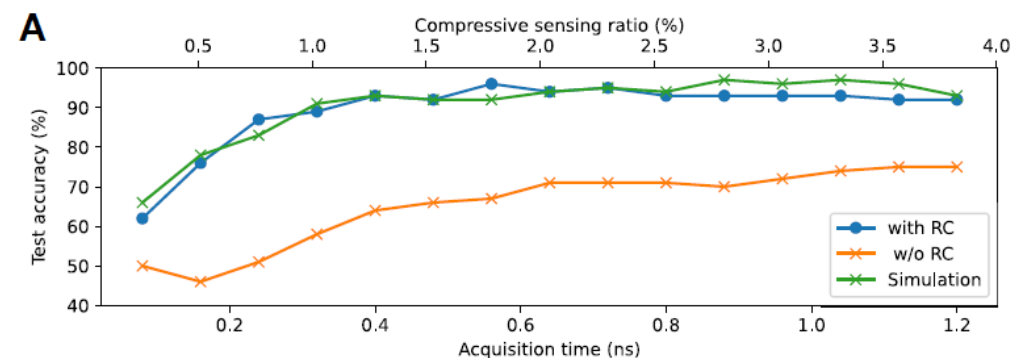
図3Aは、様々な取得時間に対する分布精度を示す。データ取得の圧縮センシング比 $C$ は $N/(28 \times 28)$ で定義され、 $N = T_N/\tau_S$ は取得波形のデータポイント数を示す。

その結果、 $T_N = 0.4 \text{ ns}$  ( $C = 1.28\%$ )で90%を超える精度が得られ、提案手法がサブナノ秒の超高速画像認識において、大幅な圧縮効率を実現できる可能性が示された。

例として、 $T_N = 0.56 \text{ ns}$  ( $C = 1.78\%$ )の混同行列を図3Bに示す。ほとんどの予測ラベルは対角線上に分布し、真のラベルと一致した。

比較のため、数値シミュレーションを行った。桁画像のランダム投影を模倣するため、 $N \times 784$ のガウスランダムマスク行列を用いた。分類器として、単一完全連結隠れ層とtanh活性化関数を持つニューラルネットワークを用いた。提案システムの分類性能は、数値ニューラルネットワークと同等であることを確認した。

# 画像認識



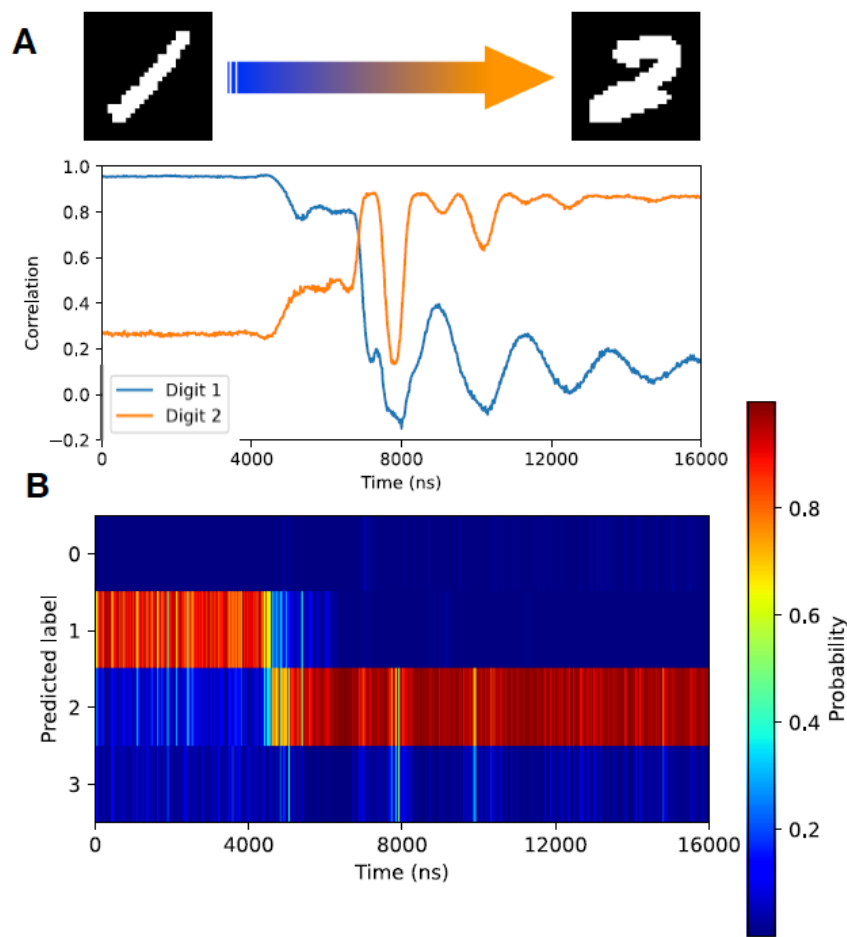
フォトニックRCプロセッサの効果を知るために、RCプロセッサに送る前のタイムドメイン信号を直接線形分類器の入力として使用し、RCプロセッサなしのシステムの分類性能を調べた。分類性能は、RCプロセッサを使用したシステムよりもはるかに悪く(図3C)、RCプロセッサからの時空間出力が分類に重要な役割を果たしていることが示唆された。

さらに、より大規模で困難な画像データセットに対する分類性能も評価した。ナノ秒の撮影時間で圧縮効率の高いデータセットでも画像分類に成功した。(補論:図S4)





# マイクロ秒現象の認識

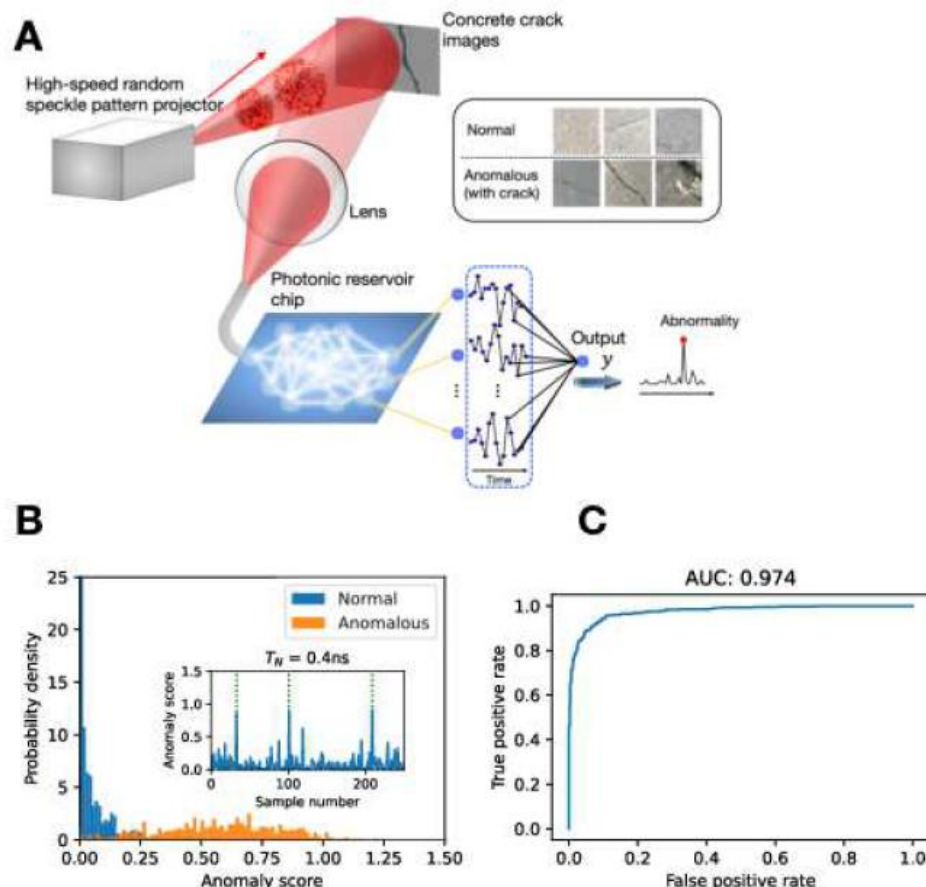


動的なシーンを認識する能力を実証するため、DMDに表示された画像が1から2に切り替わる動作を測定した。実験では、レーザー光を同じ疑似ランダム信号で繰り返し位相変調し、ダイナミックスペックルパターンを繰り返しDMDに投影した。

反射光はRCプロセッサに入力され、取得時間 $T_N = 0.56 \text{ ns}$ の間にリザーバー出力が取得され、分類結果が得られた。相関分析によると数字1から2に画像が切り替わる過渡現象は4600 nsから発生した(図4A)。

図4Bは、この過渡挙動に対する分類確率の時間依存性を示している。その結果、4600 ns付近で1から2に切り替わり、その後2を定常的に認識できることが分かった。このスイッチング動作はの検出は、相関解析の結果と一致する。DMDの表示切替の時間スケールは、数マイクロ秒に限られるが、本システムはより高速な現象認識、検出が可能である。

# イメージフリーの異常検知

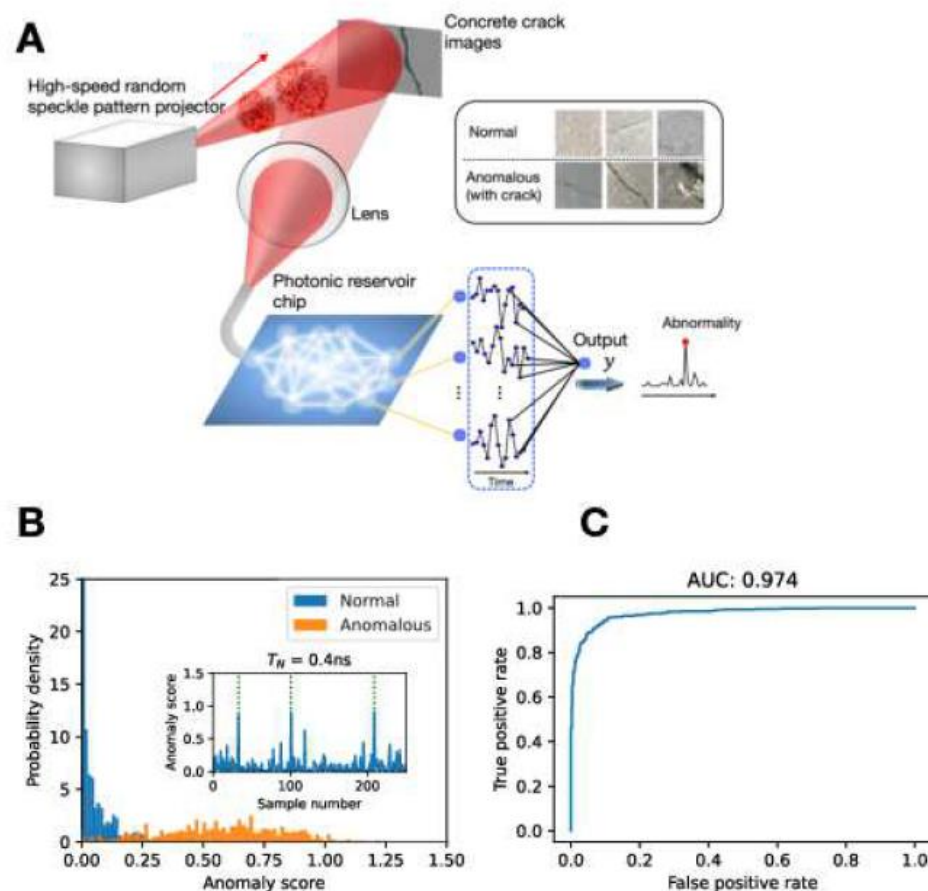


異常検出の実現可能性を評価した。異常検知は、サンプリングされた情報から異常や稀な事象を特定するアプローチであり、可能な限りリアルタイムに動作することが求められる。画像からの異常検知は一般的に計算量が多く、リアルタイムに動作の高速化を妨げる。この問題は、計算資源が限られたエッジデバイスへの実装を考えると、より深刻になる。

本アプローチでは、画像データを時系列データに圧縮変換することで、余剰で不要な情報を削減し、検出のために必要な計算をオフロードすることができる。また、画像データを他のセンサの時系列データと同様に扱えるというメリットもある。

本アプローチの特徴である軽量の計算と低い学習コストにより、エッジデバイスにおけるオンデバイス予測だけでなく、オンデバイス学習も可能になる。

# イメージフリーの異常検知

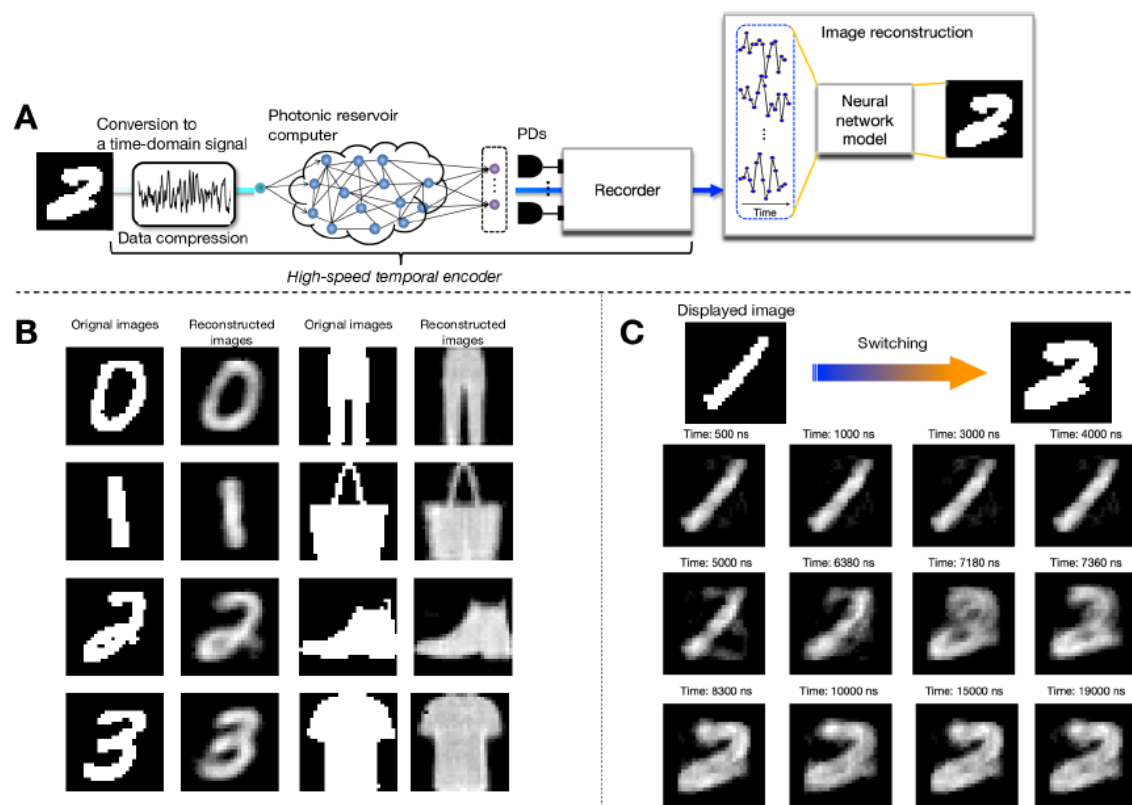


デモンストレーションのために、構造物のヘルスマニタリングと検査用のコンクリートのひび割れのベンチマークデータセットを使用した。このデータセットには、 $227 \times 227$ ピクセルのひび割れの有無が異なるコンクリート画像が含まれており、各画像は表面から約1m離れた位置で、ターゲットに直接カメラを向けて撮影された。画像はDMD上に表示された。

システムは出力 $y$ が一定値 $\alpha = 1$ に対応するように、1500の正常な(ひび無し)画像サンプルで学習された。異常(ひび有り)を測定するために、異常スコアを表現誤差 $(y - \alpha)^2$ として定義した。このスコアは正常な画像ではゼロ付近に分布するが、ひび割れが検出されると大きな異常値をもつ(図5B)。

受信者動作特性(ROC)曲線は、偽陽性率に対する真陽性率でプロットされる(図5C)。曲線下面積(AUC)は0.978であり、理想的なモデルの $AUC = 1$ を考慮すると、良好な分離性を示している。

# 画像再構成のための高速時間画像エンコーダ



ここでは、高速イメージャーとしても使用できることを示す(図6A)。重要な点は、リザーバ出力に画像情報が含まれていることである。訓練データセットで訓練したニューラルネットワークモデルを使用して、測定されたリザーバー出力からターゲットの画像を再構成した。

簡単な概念実証実験として、4クラスの手書き数字画像と4クラスのFashion-MNIST画像を含む2つのオリジナルデータセットを使用した。各画像は2値化され、DMD上に表示され、リザーバ出力は $T_N = 20 \text{ ns}$ の間に記録された。

画像再構築には、ネットワーク出力がターゲット画像に対応するように訓練された畳み込みニューラルネットワークモデルを使用した。トレーニングには900枚画像を使用し、テストには100枚の画像を使用した。

# 画像再構成のための高速時間画像エンコーダ

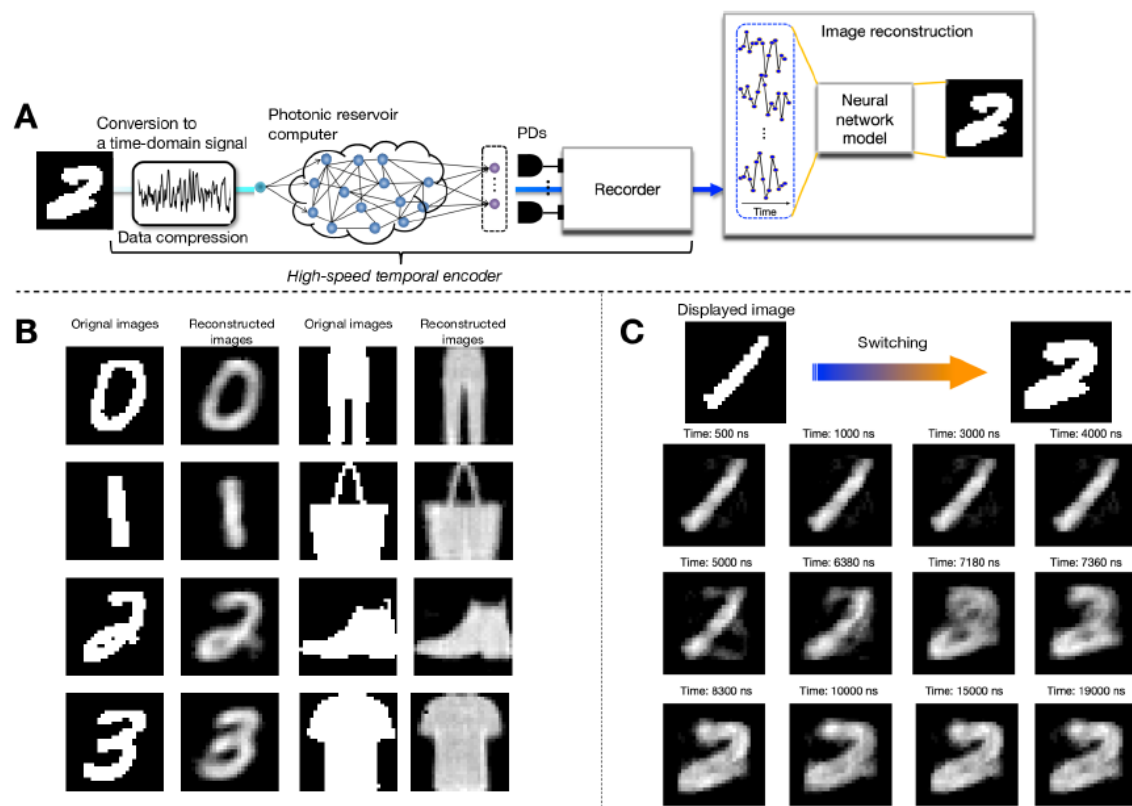
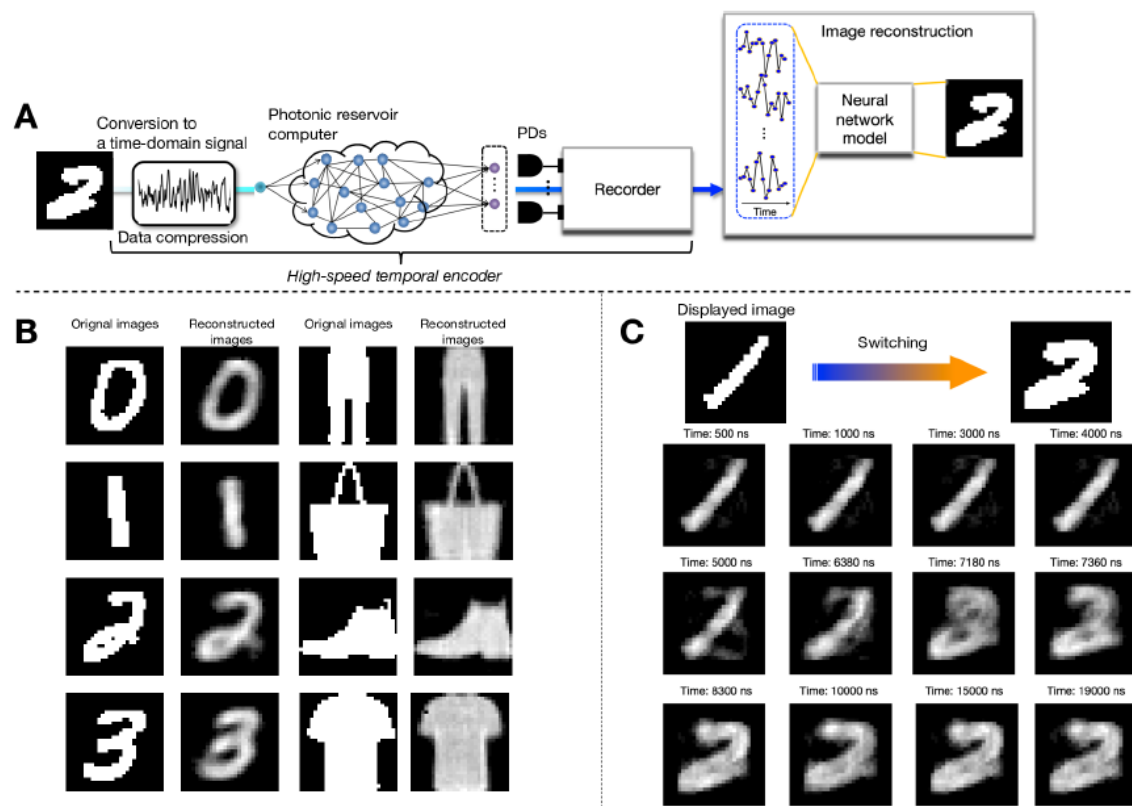


図6Bは、いくつかのテストサンプルに対する画像の再構成の結果を示している。100枚のテスト画像に対して計算された平均二乗誤差(RMSE)は、MNIST手書き数字データセットとFashion-MNISTデータセットに対して、それぞれ0.219と0.223であった。

より短い $T_N$ 値を設定すると、RMSEが増加した。しかし、このトレードオフは、波長分割多重(WDM)を組み込むことで解決できる(図S5A)。WDM方式の $T_N = 0.8 \text{ ns}$ でも同様の性能が得られた(図S6)。



# 画像再構成のための高速時間画像エンコーダ



提案する時間画像エンコーダは、稀な事象や過渡現象の観察に有効である。提案されたアプローチは、ターゲット画像の時間エンコードに広帯域パルスレーザを必要としないため、パルス幅による記録時間の制限がない。制御可能な時間分解能 $T_N$ で長時間の連続記録が可能である。

原始的な実験として実現可能性を示すため、DMDに表示された数字画像の切り替え動作を測定した。この実験では、ダイナミックスペックルパターンを繰り返しDMDに投影し、 $T_N = 20 \text{ ns}$ でリザーバー出力を取得した。この条件では、 $T_N$ の時間分解能で各時刻の画像を再構築できる。図6Cに示すように数字1から2への切り替えが観察できる。しかし、本研究では4クラスの数字画像のみを用いてネットワークを学習させたことを考慮

すると、再構成された過渡画像(図6C中段)は正しくとらえられていない可能性があり、この画像は学習された数字画像の単なる投影に帰着する。例えばハダマード基底パターンのような、独立した基底画像を含むより大きなデータセットを用いて再構築モデルを訓練することで、より正確で優れた画像化を達成することができる。

# 展望

提案されたアプローチは、ノイズに強い、強度が極めて弱い場合の画像処理などを利用できる可能性がある。  
処理速度を上げるためには、

- ・光変調器をより広帯域のものにする
  - ・空間分割多重やWDMのような様々な多重化技術を用いた光子の別の自由度に基づく並列処理をすること
    - ・空間分割多重では、複数のファイバー受信機を使用して、提案されたアプローチに組み込むことができる。
    - ・WDMでは、多波長レーザーを用いることで、独立したスペックルパターンを並列に生成することができる。
- が考えられる。

重要な課題として、エンドツーエンドの高速フォトニックプロセッサのポストプロセッサを開発することである。この技術は、平衡マッハツェンダー変調器と積分器に基づいており、積和演算を時間領域で実現することができる。時間領域でのアナログ計算のさらなる利点は、超低エネルギーで実行でき、原理的に**単一光子レベルの微弱信号を処理**できることである。エンドツーエンドのフォトニックプロセッサは、異常検出、フローサイトメモリ、ダイナミックターゲットの追跡と制御など、リアルタイム処理を必要とする幅広いアプリケーションに使用され、潜在的には未知の超高速現象の認識と検出にも使用されるであろう。