


Classificação e previsão de dados temporais usando um sistema de computação de reservatório baseado em memristor

John¹Moon,² Wen^{Ma1,2}, Jong Hoon^{Shin1}, Fuxi Cai¹,  Chao Du¹, Seung Hwan  Lee¹ e Wei D.

Lu1*

A análise de séries temporais, incluindo previsão, é essencial em uma variedade de campos, desde finanças até engenharia. No entanto, a previsão de longo prazo é difícil, particularmente para casos em que os modelos e parâmetros subjacentes são complexos e desconhecidos. Redes neurais podem processar características efetivamente em unidades temporais e são atraentes para tais fins. A computação de reservatórios, em particular, pode oferecer um processamento temporal eficiente de redes neurais recorrentes com baixo custo de treinamento e, portanto, é bem adequada para a análise de séries temporais e tarefas de previsão. Aqui, relatamos um sistema de hardware de computação de reservatório baseado em memristores de óxido de tungstênio dinâmico (WOX) que podem processar dados temporais. Os efeitos internos de memória de curto prazo dos memristores WOX permitem que o reservatório baseado em memristor não mapeie quase entradas temporais em estados de reservatório, onde as características projetadas podem ser prontamente processadas por uma função de readout linear. Usamos o sistema para demonstrar experimentalmente duas tarefas padrão de benchmarking: reconhecimento isolado de dígito falado com entradas parciais e previsão caótica do sistema. Uma alta precisão de classificação de 99,2% é obtida para a

reconhecimento de fala5 (como reconhecimento de fonema6).

oferecem uma capacidade muito melhorada de processar

Rdados temporais em comparação com redes neurais convencionais. Devido ao conectivo cíclico...

entre neurônios ocultos, ausentes das redes neurais de alimentação, as saídas nas RNNs dependem tanto dos insumos atuais quanto dos estados anteriores dos neurônios, permitindo que as RNNs descubram correlações temporais nos dados. No entanto, as conexões cíclicas em RNNs causam problemas comumente conhecidos como gradiente de desaparecimento e gradiente explosivo, o que torna o processo de treinamento caro e difícil.

Foram propostas variações de RNNs para resolver esses problemas, incluindo memória de curto prazo (LSTM)³ e computação de reservatório (RC)⁴. Em RC, em particular, um "reservatório" dinâmico que oferece uma memória de curto prazo (ou seja, memória desbotada) é usado para mapear quase as entradas temporais em um espaço de características de alta dimensão, representado pelos estados dos nós formando o reservatório. Esse mapeamento não linear pode fazer com que as entradas complexas iniciais se tornem linearmente separáveis no novo espaço com base em estados de reservatório, de modo que um processamento adicional possa ser realizado usando uma camada de rede simples e linear (muitas vezes chamada de camada de leitura)⁴. Os sistemas RC baseados em software já alcançaram desempenho de última geração para tarefas, incluindo

reconhecimento de fala5 (como reconhecimento de fonema6). Mais importante, a análise de dados temporais, que é difícil de executar para redes neurais convencionais, é naturalmente adequada para sistemas RC devido à capacidade do reservatório de mapear características diversas em diferentes escalas de tempo. Por exemplo, a previsão caótica do sistema é um problema especialmente difícil devido à alta sensibilidade do sistema ao erro, e a previsão de grandes sistemas espúriamente caóticos só recentemente foi demonstrada usando um sistema RC baseado em software^{7,8}. Os sistemas RC também mostraram desempenho superior às redes neurais convencionais em outras tarefas de previsão de séries temporais, incluindo previsão de sistemas financeiros⁹ e entrada de água¹⁰.

Estudos recentes têm como objetivo a implementação de hardware de sistemas RC, usando memristores¹¹ dinâmicos, redes de switchatróico¹², fotônicos de silício¹³ e osciladores spintrônicos¹⁴. Das diferentes abordagens, os sistemas RC baseados em memristor podem ser fabricados usando processos e materiais de fundição padrão para permitir a integração direta com elementos de controle e sensoriamento com alta densidade, tornando-os particularmente atraentes para implementações de hardware. Os memristores¹⁵⁻¹⁸ já estudaram extensivamente para aplicações de computação neuromórfica¹⁹⁻²², e estudos sobre sistemas RC baseados em memristor¹¹ mostram que as características não lineares intrínsecas e os efeitos de memória de curto prazo dos memristores fornecem propriedades centrais de um bom reservatório, namely, separação e propriedades estatísticas eco4.

Tarefas como o reconhecimento de dígitos escritos à mão foram

¹ Departamento de Engenharia Elétrica e Ciência da Computação, Universidade de Michigan, Ann Arbor, MI, EUA. ² Esses autores contribuíram igualmente: John Moon,

Wen Ma. *e-mail: wluue@umich.edu

demonstradas experimentalmente usando sistemas RC baseados em memristor¹¹.

Uma vez que o desempenho do sistema RC depende fortemente da dimensionalidade do espaço reservatório, o aumento do número de núdulos no reservatório tornou-se um desafio experimental. Uma abordagem interessante para este problema é construir o reservatório usando um único nó físico não linear submetido a feedback atrasado, que pode efetivamente acarretar como uma cadeia de nós virtuais sem muita degradação de desempenho em comparação com um reservatório convencional²³. Uma vez que apenas um único nó físico é necessário, a abordagem RC baseada em sistema de atraso é atraente para a implementação de hardware usando dispositivos emergentes, e foi demonstrada recentemente usando sistemas fotônicos e baseados em spintronic²⁴⁻²⁶.

Neste artigo, mostramos que os sistemas RC baseados em memristor que empregam o conceito de nó virtual podem ser usados para processar dados temporais de forma eficiente, produzindo excelentes resultados para tarefas importantes, como reconhecimento de fala e previsão de séries de tempo. Para exames, a previsão de séries temporais é uma área de pesquisa ativa que pode impactar campos amplos. A abordagem tradicional para a previsão de séries temporais é tentar derivar um conjunto de equações descrevendo o sistema desejado a partir de observações abundantes. No entanto, na maioria dos casos é praticamente impossível obter as equações que representam o sistema com precisão. Alternativamente, a previsão pode ser realizada com base em dados passados e atuais utilizando técnicas

Mostramos que a previsão de séries de tempo, incluindo reconhecimento de dígito falado com base em entradas parciais e previsão autônoma de sistemas caóticos, pode ser realizada com sucesso usando um sistema de hardware RC baseado em memristor. As características nas entradas temporais são efetivamente capturadas pela dinâmica nativa de curto prazo dos dispositivos memristor, em vez de loops em sistemas RC convencional^{4,7} ou feedback atrasado²³. Ao processar as características temporais mapeadas no reservatório baseado em memristor, a análise eficiente e até mesmo a previsão tornam-se viáveis. Por exemplo, em precisão de reconhecimento de fala mais de 57,8% pode ser alcançada usando apenas um quarto da sequência completa de entrada, e mais de 98,2% de precisão pode ser alcançada com 62,5% da entrada completa. A previsão de um sistema caótico, a série temporal Mackey-Glass, também é implementada experimentalmente sem saber qualquer informação sobre a equação descrevendo o sistema. A previsão autônoma de até 50 etapas é demonstrada aplicando o próximo passo previsto como entrada no sistema sem qualquer feedback externo, e uma previsão muito mais longa se torna possível através de atualizações periódicas que impedem o reservatório baseado em memristor de divergir da dinâmica original.

Sistema RC baseado em memristor

Em um memristor típico, o estímulo elétrico aplicado desencadeia a migração de vagas de oxigênio ou íons metálicos no layer de comutação, levando a modulações da resistividade local e da

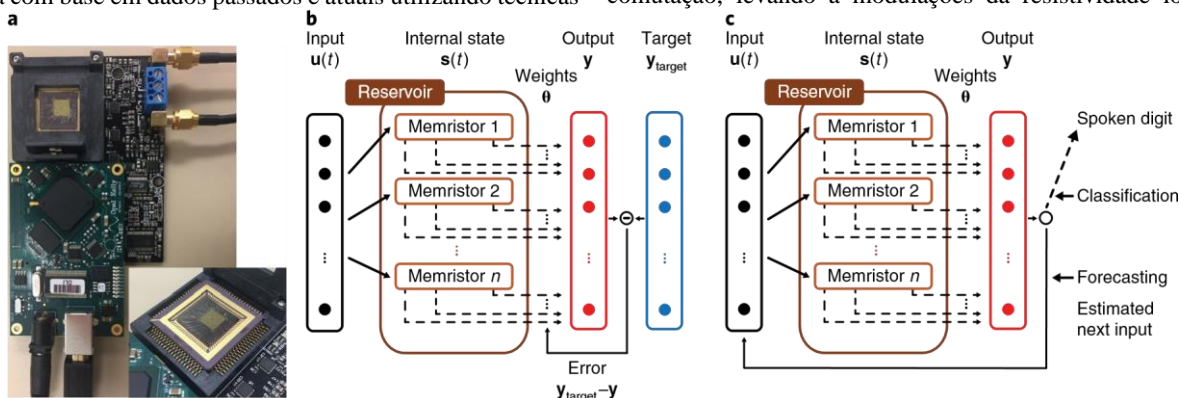


Fig. 1 | Sistema RC baseado em memristor. **a**, imagem óptica do sistema de hardware. A matriz memristor é ligada a fios e montada no sistema de placa de teste. Inset: image óptico do chip 32×32 memristor. **b**, Esquema da fase de treinamento para o sistema RC baseado em memristor. Os pesos na camada de leitura são atualizados para reduzir o erro de saída. **c**, Esquema da fase de teste para o sistema RC baseado em memristor. O sistema pode ser usado para executar tarefas de classificação ou prever tarefas onde a saída produzida a partir do sistema é devolvida ao reservatório como entrada para o próximo quadro.

de aprendizagem de máquina baseadas em estatísticas^{27,28} ou técnicas de aprendizado de máquina^{29,30}. No entanto, a precisão de previsão das técnicas com base em estatísticas é limitada pelos parâmetros do modelo assumido e pela complexidade do sistema. Sistemas típicos muitas vezes não conseguem capturar adequadamente as relationships não lineares nos dados, mesmo com dependências não lineares incluídas no modelo. As redes neurais oferecem uma ferramenta mais geral e flexível, pois não dependem de parâmetros de tarefas específicas, mas são impulsionadas apenas pelos dados³¹. Em particular, as RNNs têm reunido muita atenção para a previsão, uma vez que as informações temporais capturadas pelas conexões recorrentes melhoram o desempenho da previsão³²⁻³⁴, embora a um custo de treinamento aumentado. Para isso, os sistemas RC oferecem a capacidade de processamento temporal eficiente com baixo custo de treinamento e, portanto, são adequados para tarefas de análise e previsão de séries temporais, incluindo previsão autônoma de sistemas caóticos^{7,8,26}.

resistência geral do dispositivo³⁵. Em alguns dispositivos, o sistema pode relaxar rapidamente de volta ao estado original devido à difusão espontânea de íons, levando a um comportamento de memória de curto prazo^{36,37}. Esses dispositivos oferecem nativamente as propriedades de 'memória desbotada' exigidas por um reservatório e têm sido usados para implementar sistemas RC para tarefas como reconhecimento de dígitos escritos à mão, onde os recursos no domínio espacial são convertidos em recursos em entradas de streaming¹¹.

Observamos, no entanto, que há grandes diferenças nas configurações dos reservatórios entre um sistema RC convencional e o baseado em memristor. Os reservatórios típicos podem consistir de várias centenas ou milhares de nós internos através de interconexões complexas, onde os estados dos nós são todos acessíveis pela rede de leitura. Na implementação baseada em memristor, o efeito de memória de curto prazo é obtido nativamente a partir de um único dispositivo, em vez de loops formados por múltiplos núdulos. Como resultado, um

reservatório pode consistir em um único dispositivo (nó) onde o estado do reservatório é representado pelo estado de memristor animado e extraído de medições de condução. Isso permite uma implementação experimental muito mais simples, mas também limita o tamanho do reservatório. Os dispositivos Multiple podem ser usados para expandir o tamanho do reservatório com base em variações de dispositivo para dispositivo onde o estado do reservatório é representado pelos estados coletivos de todos os dispositivos¹¹¹, mas os nós são independentes uns dos outros neste caso, em vez de serem quase acoplados. Para expandir a dimensionalidade do reservatório e melhorar o desempenho do sistema, adaptamos o conceito de nós virtuais originalmente desenvolvidos em sistemas de atraso²³. Em um sistema de atraso, um único nó físico não linear é submetido a feedback delayed, onde as excitações do nó físico em resposta aos sinais atrasados podem efetivamente agir como uma cadeia de nós virtuais²³. O estado dos nós virtuais depende do próprio estado anterior do nó, do estado atual dos nós adjacentes e do sinal de entrada mascarado, permitindo que eles não sejam quase acoplados. Ao usar máscaras geradas aleatoriamente para sinais de entrada, diversas respostas podem ser obtidas a partir dos nós virtuais, e os sistemas de atraso têm se mostrado capazes de alcançar comparação de desempenho com a dos reservatórios convencionais e bem projetados²³.

Nós temos a hipótese de que o conceito geral de nó virtual, representado como a resposta não linear de um dispositivo memristor físico em etapas de tempo selecionadas em resposta a uma entrada de streaming, pode ser usado da mesma forma para aumentar o tamanho do reservatório e permitir um melhor mapeamento dos recursos de entrada. Abaixo implementamos experimentalmente esse conceito e mostramos que essa abordagem de fato permite que o sistema RC baseado em memristor execute com sucesso tarefas complexas, como análise e previsão de dados temporal.

A implementação do hardware é baseada em uma matriz de memristor de 32×32 WOX (Figura Suplementar 1 e Note 1), montada em uma placa de teste personalizada (Fig. 1a e Métodos). O funcionamento do sistema RC é mostrado em Fig. 1b,c, para a fase de treinamento e a fase de testes, respectivamente. Durante a fase de treinamento, um sinal de professor $u(t)$ é aplicado ao reservatório para levar o reservatório para diferentes estados animados (dependendo dos insumos de fluxo temporal). Os pesos ϕ na camada de leitura são então treinados para minimizar o erro entre a saída da rede de leitura y e o alvo de saída de destino. O alvo pode ser o rótulo de dígito correto (para classificação) ou o próximo valor nos dados da série temporal (para previsão). Note que apenas pesos na camada de leitura linear precisam ser treinados, enquanto as conexões no reservatório permanecem fixas. Na fase de testes, os insumos de teste, não incluídos no conjunto de treinamento, são aplicados ao reservatório. Dependendo do tipo de tarefa, a implementação final da saída pode ser bem diferente. Para classificação, a configuração do sistema permanece a mesma da fase de treinamento, e o desempenho do sistema é avaliado analisando quantas amostras de teste são corretamente classificadas. Por outro lado, na previsão de aplicações, a saída da camada de leitura é devolvida ao sistema como sinal de entrada para o reservatório na próxima etapa. Neste caso, as entradas externas não são necessárias, e o sistema é deixado em execução autônoma para gerar a série de tempo de destino.

Reconhecimento de fala

Para validar a abordagem proposta, primeiro realizamos um benchmark padrão de reconhecimento de voz — reconhecimento

isolado de dígito falado usando o sistema RC baseado em memristor. Os insumos para o reservatório são formas de onda sonora de dígitos falados isolados (0-9 em inglês) from o banco de dados NIST T14638, pré-processado usando o modelo de ouvido passivo de Lyon³⁹ baseado em canais cocleares humanos. O pré-processamento transforma as formas de onda sonora em um conjunto de vetores de 50 dimensões (correspondentes aos canais de frequência) com até 40 etapas temporais (Figura Suplementar 2 e Nota 2). Um exemplo representando o dígito 0 é mostrado na Fig. 2a. Cada ponto de dados no gráfico representa a probabilidade de disparo de um neurônio correspondente a um canal de frequência específico em um ponto de tempo, com 50 channels de frequência no total plotados ao longo do eixo y e tempo traçado ao longo do eixo x . O gráfico de entrada (chamado de cochleograma) é então digitalizado para formar as entradas de fluxo para o reservatório⁴⁰ (Fig. 2b). Para esta tarefa de classificação de dígito falado isolado, 50 dispositivos memristor WOX são usados para implementar experimentalmente o sistema RC, com cada dispositivo processando o trem de pico de entrada em um dos 50 canais (ver Métodos).

A Figura 2c mostra as respostas obtidas experimentalmente dos memristors para os trens de pico nos canais de frequência 47 e 19, respectivamente. Os dados de entrada foram extraídos do orador feminino 1, primeiro pronunciamento no banco de dados. Devido à dinâmica de curto prazo dos memristors, os padrões temporais nos trens de pico de entrada levaram a diversos, mas determinísticos responses de dispositivos, onde a condutância do dispositivo é aumentada quando estimulada por um pico e decai espontaneamente, com a condução do dispositivo em um momento específico dependendo do histórico recente dos insumos temporais¹¹. No entanto, as informações temporais da parte inicial da sequência de entrada (isto é, muito história) não são transmitidas nas respostas finais dos memristors. Por exemplo, como mostrado na Fig. 2c, embora as respostas do dispositivo sejam claramente diferentes dentro das etapas de tempo 10-35 das entradas de streaming, as respostas no final no momento passo 40 não mostram grande diferença entre as duas entradas devido a semelhanças na última parte das duas entradas (passos de tempo 36-40). A perda de informações levará a resultados de classificação ruins. Para resolver esse problema, dividimos toda a sequência de entrada em n intervalos iguais e medimos e registramos o estado do dispositivo no final de cada intervalo. Esse processo efetivamente cria n nódulos virtuais a partir de um único dispositivo. O estado de nó virtual extraído é afetado pelo pattern de entrada temporal na história próxima desse nó virtual, e também pelos estados de nó virtual anteriores. A cadeia de estados de nó virtual, por sua vez, forma nódulos no reservatório para permitir uma análise efetiva dos dados temporais.

Especificamente, medimos a condução do dispositivo nos passos de tempo $40/n$, $40 \times 2/n$, $40 \times 3/n$, $40 \times 4/n$, ..., $40 \times (n-1)/n$ e 40 como estados dos nódulos virtuais no reservatório. Um exemplo de $n = 8$ é mostrado em Fig. 2c, onde os pontos vermelhos representam os estados de nó virtual medidos. Os estados de nó virtual são fornecidos à camada de leitura (um perceptron de 400×10 neste caso, ver Métodos) para treinamento e teste. A Figura 2d mostra a matriz de confusão obtida experimentalmente durante os testes, após o treinamento da camada de leitura com 450 amostras de fala. No geral, uma alta taxa de reconhecimento de mais de 99,2% foi obtida para todas as entradas usando o sistema RC baseado em memristor (ver Métodos).

Para verificar a eficácia do reservatório baseado em memristor na captura de características temporais na entrada, o sistema baseado em memristor foi equipado com outros sistemas baseados

em filtros convolucionais convencionais. O sistema baseado em memristor mostra vantagens claras no mesmo tamanho da rede, devido à capacidade do dispositivo de realizar nativamente a transformação não linear da entrada, bem como a capability de capturar características temporais locais dentro de uma janela de entrada e as características mais globais entre janelas de entrada, habilitadas pelo efeito de memória interna de curto prazo dos dispositivos (Fig. 3 e Nota Suplementar 3).

Mais interessante, à medida que o reservatório baseado em memristor mapeia as características temporais associadas a diferentes classes de entradas, é possível usar o sistema RC para prever o dígito falado antes que a expressão seja concluída. Esta hipótese foi testada experimentalmente, como mostrado na Fig. 3. Quando os primeiros 25% do sinal de fala são usados, cada dispositivo produz apenas um nó virtual, de modo que uma rede de leitura de 100×10 é usada durante a fase de treinamento. Como mostrado na Fig. 3b, após o treinamento uma taxa de reconhecimento de 57,8% pode ser alcançada usando apenas os primeiros 25% de toda a sequência. A precisão de reconhecimento aumenta quando a parte do insumo disponível aumenta (por exemplo, 98,2% de taxa de reconhecimento para usar os primeiros 62,5% do insumo e 99,2% de taxa de reconhecimento quando toda a sequência de fala é usada). A evolução da taxa de reconhecimento em função da porção de entrada utilizada é mostrada na Figura Suplementar 4. A taxa de reconhecimento de 99,2% é comparável aos resultados alcançados em sistemas RC baseados em dispositivos ^{spintronic14} ou fotônico^{24,25}, embora este experimento também tenha verificado a capacidade do sistema RC de realizar a classificação de dígito falado usando apenas parte da sequência de entrada.

Previsão da série temporal

A tarefa isolada de reconhecimento de dígito falado tem relativamente boa tolerância ao erro de predição porque a tarefa requer apenas a identificação da maior saída entre as 10 saídas. Assim, enquanto a etiqueta selecionada estiver correta, pequenos erros no valor exato de saída não degradam o desempenho do sistema. Em contraste, a previsão de dados de timeseries é uma tarefa muito mais difícil, uma vez que a diferença quantitativa entre a verdade do solo e o valor previsto importa e a diferença pode ser ainda maior acumulada em previsões subsequentes. Um teste padrão de referência para a previsão de séries temporais é prever um sistema caótico, que é inerentemente muito desafiador devido ao expoente lyapunov ^{positivo41} em sistemas caóticos, o que leva a um crescimento exponencial da separação de trajetórias próximas para que mesmo pequenos erros de previsão possam levar a uma divergência da previsão da verdade terrestre.

Para ver se o sistema RC baseado em memristor pode ser usado para previsão de longo prazo, testamos o sistema usando o tempo Mackey-Glass ~~tem uma~~ forma determinística, mas são difíceis de

$1\beta\delta x \text{ t}\delta \pi\phi\beta n \gamma x\delta t\beta$. Estes tipos de sistema ^{caótico7,26} e, portanto, ^{série42,43} $\frac{ddx}{dt} = 1/4 - \beta x\delta t - \beta$

têm sido amplamente utilizados como referência para a previsão de testes de tarefas. Para obter dinâmica caótica, estabelecemos os parâmetros $\beta = 0,2$, $\gamma = 0,1$, $\tau = 18$, $n = 10$ em nossos estudos (ver Métodos).

Para melhorar a precisão da previsão, é necessário construir um sistema de reservatórios que possa capturar a dinâmica temporal do determinado sistema caótico o mais próximo possível, o que, por sua vez, permite que a rede de leitura alcance uma previsão precisa. Para alcançar esse objetivo, utilizamos várias técnicas para ampliar a dinâmica do reservatório. First, usamos memristors diferentes para o nosso sistema de reservatórios baseado em memristor. Devido às variações de dispositivo para dispositivo

(Fig. 5 Suplementar), os dispositivos produzem respostas qualitativamente semelhantes, mas quantitativamente diferentes, mesmo com a mesma entrada, e tais variações de dispositivos expandem a resposta do reservatório. Em segundo lugar, n os nós virtuais são obtidos de cada memristor físico, com base na resposta do dispositivo na etapa atual e $n - 1$ etapas temporais anteriores, semelhante à abordagem utilizada na tarefa de classificação (Fig. 6 suplementar e Nota 4).

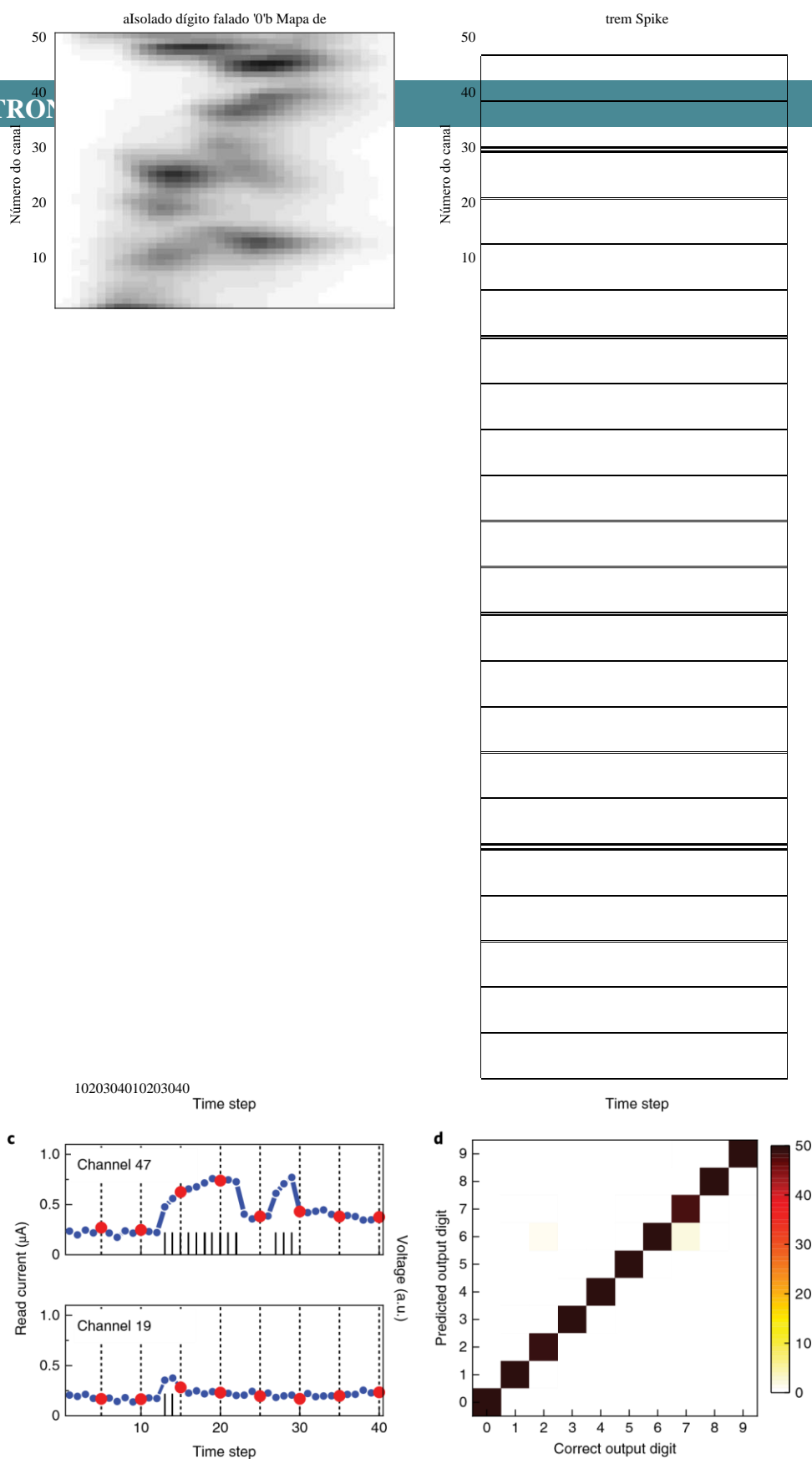
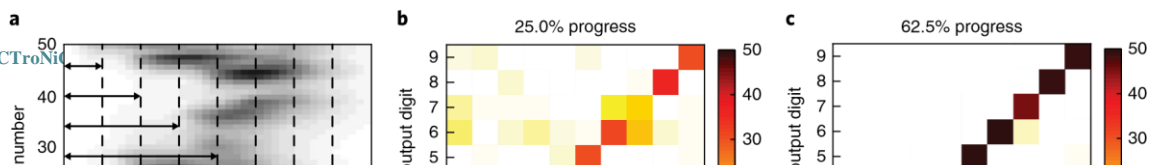


Fig. 2 | Implementação da tarefa de reconhecimento de dígito falado. a, Cochleagram da primeira palestrante feminina, primeira amostra de fala de expressão após ser p... ponto de dados representa a probabilidade de disparo de uma célula ciliada sensível a uma determinada frequência (canal) em um determinado ponto de tempo. **b**, trens de es... mostrado em **a**, definindo um limiar (0,5) para a probabilidade de disparo. **c**, Resposta temporal dos memristors aos trens de pico nos canais 47 e 19. A corrente (pontos azu... As entradas de tensão do trem de espigão (linhas pretas) também são traçadas para referência. A entrada completa pode ser dividida em oito intervalos, representando oito no... (pontos vermelhos). **d**, Matriz de confusão mostrando os resultados de classificação obtidos experimentalmente do sistema de reservatório baseado em memristor versus as s... é alcançada. Barra de cor: ocorrência de uma determinada saída prevista.



A capacidade do sistema RC baseado em memristor na previsão de séries temporais foi testada tanto experimentalmente usando um reservatório com $m = 20$ dispositivos e $n = 50$ nós virtuais para cada dispositivo, quanto através de simulações do reservatório usando um modelo de dispositivo realista (ver Métodos). Os estados do reservatório são então aplicados à camada de leitura (uma rede de 1.000×1 , ver Métodos) para gerar os dados previstos para a próxima etapa. A Figura 4a,c mostra os resultados obtidos durante o treino das medições experimentais e da simulação, respectivamente. Excelente concordância entre a meta e o valor preducido pode ser obtido, indicando que a leitura treinada pesos podem calcular corretamente o próximo sinal de passo-de-tempo com base nos estados internos do reservatório. Outras evidências de treinamento bem-sucedido podem ser encontradas examinando o desempenho da rede no domínio de frequência e no espaço de fase, como mostrado na Fig. 4e,f para os resultados experimentais, onde uma excelente correspondência pode ser novamente observada entre a saída memristor RC e a verdade do solo. Os resultados experimentais são ainda mais verificados através de simulação, como mostrado na Fig. 4c e Figura Suplementar 7.

A rede é então usada para prever a série temporal de forma autônoma. Antes do início da previsão autônoma, um

previsão, plotado no domínio de frequência (g) e no espaço de fase (h).

o estágio de inicialização é necessário para preparar os estados internos do reservatório, uma vez que os sistemas caóticos dependem fortemente das condições iniciais. Durante a fase de inicialização, a conexão de feedback entre a saída da rede e a entrada do reservatório é removida, e os dados de entrada verdadeiros são anexados ao reservatório. Note-se que o objetivo da inicialização é simplesmente excitar o reservatório para o estado pouco antes do início da previsão autônoma, não treinar o sistema. Após a inicialização do reservatório, a saída do function de leitura, ou seja, os dados previstos para o próximo passo, é então conectada ao reservatório como nova entrada, e o sistema produz continuamente a série temporal prevista.

A Figura 4b,d mostra os resultados das medições experimentais e simulação, respectivamente, para a previsão de séries temporais autônomas usando o sistema RC baseado em memristor. À medida que os estados do reservatório são estabilizados durante a fase de inicialização, os valores de saída obtidos da camada de leitura seguem com precisão os valores corretos. Depois disso, a saída gerada de forma autônoma (a partir do 500º passo em diante) ainda corresponde muito bem à verdade do solo (que não é usada na

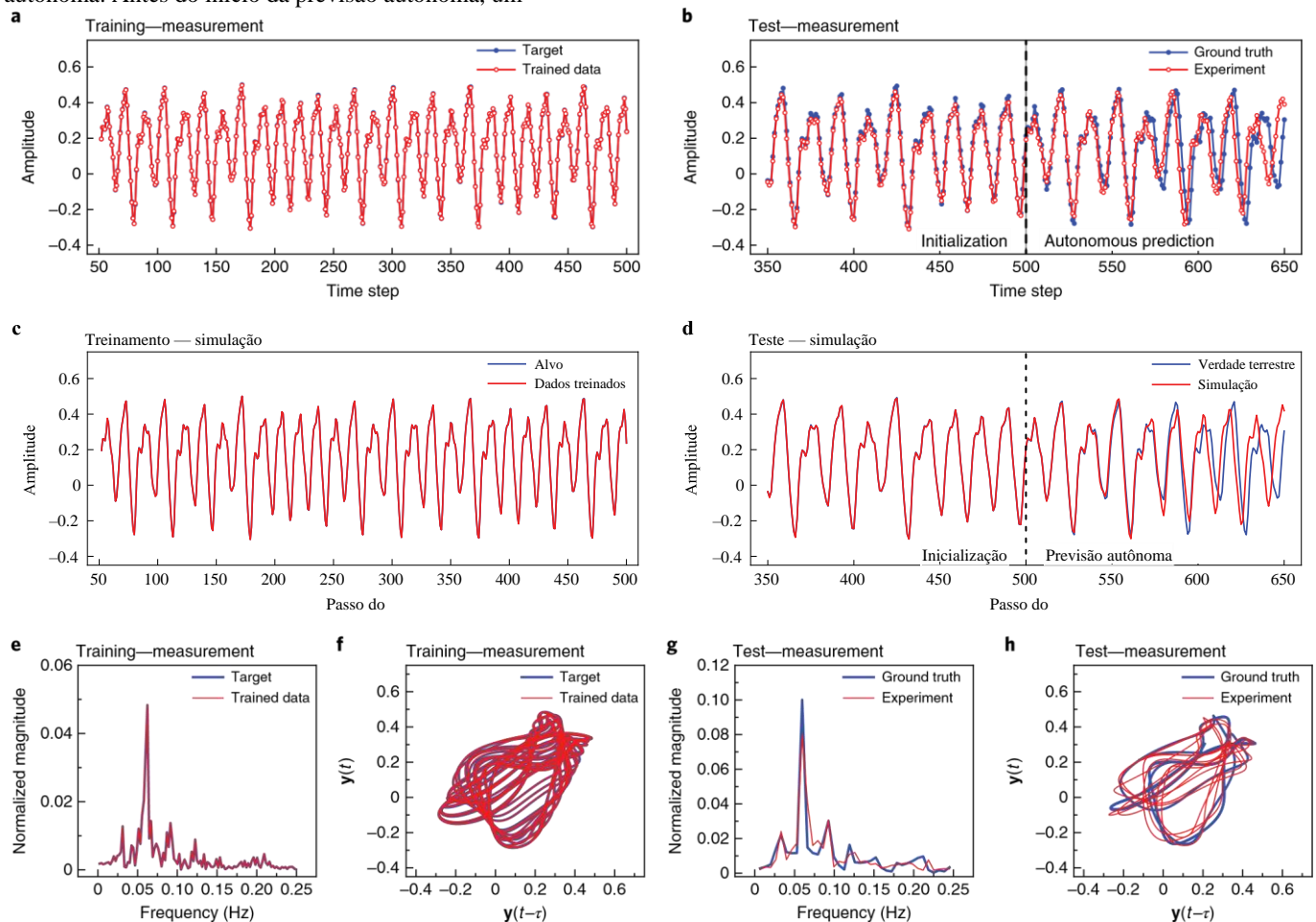


Fig. 4 | previsão autônoma da série temporal Mackey-Glass. a,b, Treinamento (a) e previsão (b) resultados obtidos experimentalmente do sistema RC memristorbased. A verdade do solo (azul) e a saída prevista do sistema RC (vermelho) são plotadas. Uma etapa de inicialização é usada antes do processo de previsão autônoma, que começa a partir do tempo step 501 em b. c,d, Resultados de treinamento (c) e previsão (d) obtidos a partir da simulação usando um modelo de dispositivo realista e a mesma configuração do sistema RC como nos experimentos. e,f, Resultados experimentais da saída do sistema RC memristor plotados na frequência domain (e) e no espaço de fase (f) para a fase de treinamento. g,h, Resultados experimentais da saída do sistema RC memristor durante a produção autônoma

geração de sinal, mas simplesmente usada como referência na trama), mostrando a capacidade do sistema RC baseado em memristor de prever autônomo o sistema caótico. Após 60-70 passos tempos de previsão autônoma, o sinal previsto começa a divergir do valor correto, tanto na simulação quanto no experimento.

Uma análise cuidadosa da divergência mostra que pequenos erros na previsão são acumulados durante a previsão autônoma, podendo levar a uma mudança de fase dos dados da série temporal representados pela falta de um ou mais pontos de dados na previsão, resultando em maior erro de previsão, como mostrado na Fig. 4b no tempo passos 570-650 (Nota Suplementar 5). No entanto, embora no longo prazo as mudanças de fase ocorram devido ao pequeno tamanho da rede no experimento e ao acúmulo de erros, o sistema RC baseado em memristor ainda pode capturar corretamente a dinâmica temporal característica da série Mackey-Glass (ver Figura Suplementar 7 para dados de até 1.000 etapas temporais e os gráficos de rastreamento em Vídeo Suplementar 1). Curiosamente, com as mudanças de fase muitas vezes observamos a previsão autônoma para preceder a verdade do solo (versus seguir a verdade terrestre), mantendo a mesma dinâmica característica (Vídeo Suplementar 1). O exame dos dados no domínio de frequência também mostra picos de frequência semelhantes nos dados gerados de forma autônoma e na verdade do solo (Fig. 4g), e traços de traços semelhantes podem ser observados no espaço de fase (Fig. 4h). Notamos que se o sistema RC baseado em memristor não pudesse emular a dinâmica do sistema caótico, a previsão convergiria para um ponto estável ou órbitas periódicas. De fato, se o tamanho do reservatório for reduzido, por exemplo, para apenas um dispositivo físico, as previsões autônomas obtidas tanto no experimento quanto na simulação decaem gradualmente

direcionados. Nesse sentido, as variações de dispositivo para dispositivo são benéficas para ajudar a expandir o espaço do reservatório e melhorar o desempenho do sistema RC.

Aumentar ainda mais o tamanho do reservatório usando mais memristors e usando mais estados anteriores (nós virtuais) pode reduzir ainda mais o erro de previsão para que o comprimento da previsão precisa possa ser aumentado ainda mais. No entanto, aumentará o custo de implementação do Hardware, bem como o custo de treinamento da camada de leitura. Além disso, mesmo com um erro reduzido usando um reservatório maior, o acúmulo do erro de previsão durante a previsão autônoma é inevitável, e mais importante o tempo de previsão precisa não aumentará proporcionalmente, uma vez que o erro geral cresce exponencialmente no tempo. Assim, a previsão autônoma divergirá exponencialmente do sistema caótico original, não importa quão pequeno seja o erro de previsão inicial.

Em vez de aumentar o tamanho da rede para estender o alcance da previsão correta ainda mais para o futuro, exploramos uma abordagem onde o reservatório prevê moderadamente para o futuro, e o estado do reservatório é então periodicamente empurrado de volta para a dinâmica original antes que a previsão diverga significativamente da verdade do solo. Esta etapa de "atualização" é semelhante à fase de inicialização, e é aplicada após um período de previsão autônoma enviando a entrada verdadeira em vez do valor previsto para o reservatório por um curto período de tempo, como mostrado na Fig. 5. Não é necessário retreinar na fase de atualização. Com essas atualizações periódicas, a previsão de longo prazo de sistemas caóticos torna-se viável. Por exemplo, implementando iterativamente 50 etapas temporais de previsão autônoma, seguida por 25 etapas de tempo de atualização, a previsão experimental mostra excelente concordância com os valores reais acima de 2.000 etapas

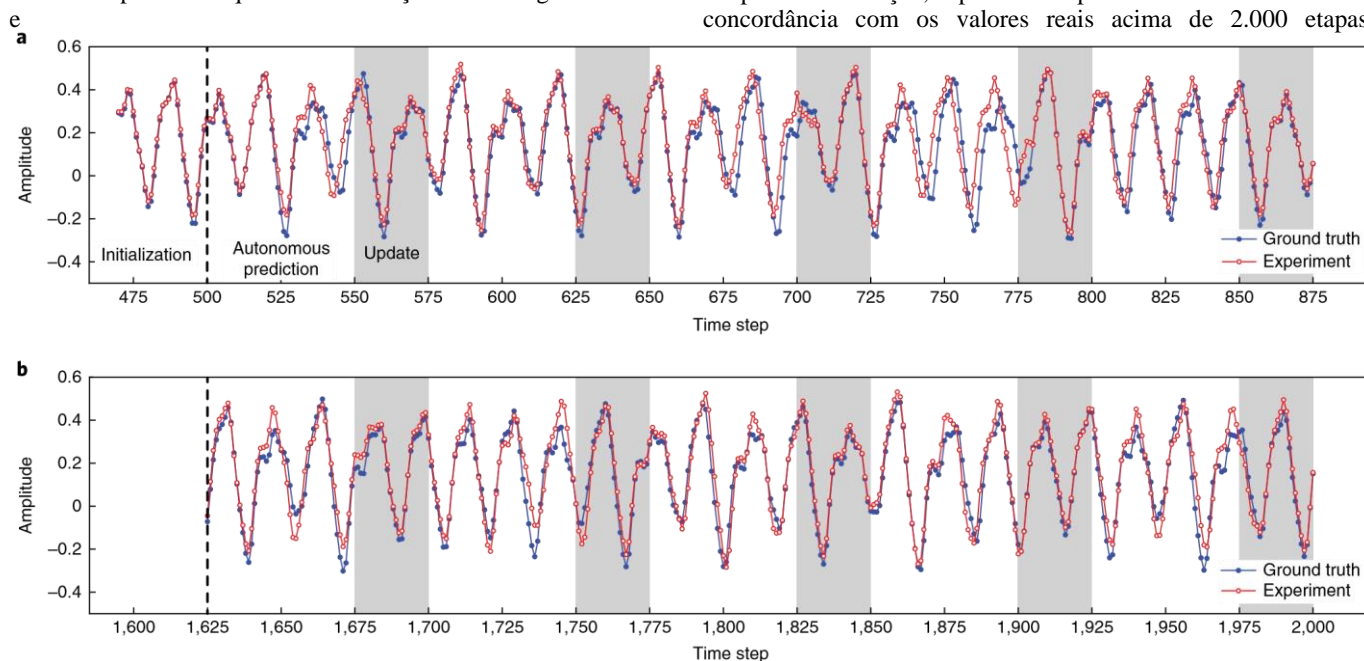


Fig. 5 | Previsão de longo prazo da série de tempo Mackey-Glass com atualizações periódicas. **a**, Saída experimental do sistema RC baseado em memristor no início do teste, mostrando resultados dos segmentos de previsão autônoma (50 passos cada), seguido pelos segmentos de atualização (regiões sombreadas, 25 passos cada). **b**, Saída experimental do sistema RC baseado em memristor no final do teste. Previsão confiável ainda pode ser obtida, auxiliada por atualizações periódicas.

finalmente convergir para um ponto estável (Supplementary Fig. 8 e Note 6) em vez de mostrar os comportamentos caóticos

temporais (limitada apenas pelo tamanho tampão da placa de teste). Durante a fase de atualização, o reservatório inverte o desvio do sistema caótico original causado pelo erro de previsão, e após a sequência de atualização o reservatório baseado em

memristor pode novamente executar a previsão para os próximos 50 passos.

Em comparação com a previsão autônoma sem atualizações, a prediction com atualizações periódicas produz resultados mais estáveis e precisos, tanto nos domínios de tempo e frequência quanto no espaço de fase (Figura Suplementar 9). Observe que, em muitos casos, a verdade do solo pode de fato ser medida periodicamente, de modo que a atualização periódica do reservatório para manter a operação estável do sistema é viável, e previsões confiáveis podem ser obtidas durante os períodos em que a verdade do solo não pode ser medida.

Direções futuras para desenvolvimentos de arquitetura e dispositivos

Notamos que a capacidade de capturar diversas características temporais em sistemas RC, incluindo o sistema RC baseado em memristor discutido aqui, é fundamental para que o sistema execute tarefas complexas de processamento de dados temporais, como a previsão caótica do sistema. Nossos estudos de controle sobre predição autônoma usando redes neurais convencionais mostram que esses sistemas ficam facilmente presos em um padrão periódico em vez de produzir o padrão caótico desejado, mesmo em tamanhos de modelo muito maiores do que o sistema RC (Figs Suplementares. 10 e 11 e Nota 7). Embora as funções de ativação do neurônio não sejam lineares em redes neurais de alimentação, os pesos e somas são lineares para que as entradas de diferentes etapas do tempo sejam processadas da mesma forma. Em contraste, os sistemas RC transformam entradas de diferentes etapas do tempo de forma muito não linear, permitindo que esses sistemas capturem diversas características temporais e dando-lhes melhor capacidade de processar dados temporais.

Em comparação com as implementações digitais de sistemas RC, a implementação de hardware baseada em memristor oferece um custo de energia muito menor (Fig. Suplementar 12, Tabela 1 e Nota 8). A energia pode ser reduzida reduzindo ainda mais a corrente de programação dos dispositivos e usando pulsos de entrada mais curtos.

Em nossa implementação atual, as variações de dispositivo para dispositivo ajudam a melhorar o desempenho do sistema, expandindo o tamanho efetivo do reservatório. No entanto, as variações aleatórias de dispositivo para dispositivo dificultam a construção de cópias dos sistemas RC treinados para tarefas de inferência. No futuro, esperamos que as características do dispositivo intencionalmente controlados que ofereçam escalas de tempo distintas sejam obtidas, controlando firmemente as condições de fabricação, como o tempo de oxidação e a temperatura para diferentes dispositivos no ^{reservatório}⁴⁴. Um melhor desempenho de RC também pode ser esperado, uma vez que os recursos podem ser capturados em escalas de tempo mais amplas não alcançáveis apenas a partir de variações aleatórias de dispositivos.

Em contraste com as variações de dispositivo para dispositivo, variações ciclo a ciclo são prejudiciais para o desempenho do sistema, uma vez que levam a erros durante a transformação do sinal temporal através do reservatório. Os dispositivos memristor baseados em WOX apresentaram boas características de ciclismo; no entanto, pequenas variações ciclo-ciclo a ciclo podem ser observadas (Fig. 13 suplementar e Nota 9). A variação ciclo-ciclo a ciclo atua como uma fonte de erro durante a implementação experimental, e pode explicar as pequenas diferenças nos resultados experimentais e de simulação obtidos para a tarefa caótica de pré-idade do sistema. Outras otimizações de dispositivos que reduzem a estocástica durante o processo de

comutação serão desejáveis para ajudar a minimizar ainda mais as variações de ciclo a ciclo e melhorar o desempenho do sistema.

A capacidade de computação de um sistema RC é, em sua maioria, determinada pelo quão bem o reservatório pode capturar as diversas características temporais e mapear essas características para o estado do reservatório. Semelhante às redes convencionais, a expansão da profundidade ou largura do reservatório pode aumentar ainda mais a capacidade de computação. Systems profundos de RC que utilizam subreservoirs em camadas podem ajudar o sistema a extrair recursos em várias escalas de ^{tempo}^{45,46}. Estudos ^{teóricos}^{45,47} também sugeriram que tais sistemas RC profundos podem operar perto da borda do caos, que é um ponto operacional ideal conhecido por RC. A largura dos sistemas RC pode ser expandida da mesma forma utilizando subreservoirs com dinâmicas em diferentes escalas de tempo (Fig. Suplementar 14 e Note 10).

Tarefas mais práticas de previsão, como previsão de longo prazo de manchas solares mensais (Fig. 15) e tendências da indústria e do consumidor (Fig. 16 Suplementar) também parecem possíveis. Para dados que mostram comportamentos limpos e periódicos, como o caso da produção de eletricidade, as ferramentas de previsão baseadas em estatísticas podem mostrar desempenho comparável ou até melhor; no entanto, algoritmos de aprendizagem de máquina, incluindo o LSTM e o sistema RC baseado em memristor mostram melhorias quando os dados são barulhentos (como o exemplo de produção de cerveja, Fig. Suplementar 17 e Note 11). Além disso, o sistema RC baseado em memristor mostra um desempenho claramente melhor quando os dados são muito complexos (como o exemplo mackey-glass, Complementar Fig. 18), enquanto previsões de ambas as ferramentas baseadas em estatísticas e LSTM ficam presas em comportamentos periódicos. Em geral, o processamento de tarefas mais complexas requer a capacidade de capturar diversas características temporais, e o sistema RC baseado em memristor pode oferecer uma vantagem nesses casos.

Conclusões

Neste estudo, um sistema RC baseado em memristor que utiliza a dinâmica iônica interna de curto prazo dos dispositivos memristor e o conceito de nódulos virtuais é demonstrado com sucesso para tarefas de análise e previsão de séries temporais. Obteve-se alta precisão de classificação de 99,2% para reconhecimento de dígito falado. Como o reservatório mapeia as características temporais do insumo, ainda foi obtida boa classificação mesmo com entradas parciais. Com o uso de atualizações periódicas para trazer o reservatório de volta à dinâmica original, a previsão pode ser mantida a longo prazo sem retrainar o sistema mesmo para tarefas caóticas.

O conjunto de crossbar memristor usado neste trabalho fornece os dispositivos de alta densidade onde os dispositivos memristor no reservatório funcionam de forma independente e paralelamente ao processamento dos dados espessotemporal, por exemplo, entradas de diferentes canais de frequência. Outras melhorias no sistema do reservatório podem envolver o uso de dispositivos interconectados para formar estruturas de reservatórios mais complexas, com pesos e loops de conexão devidamente projetados no sistema. Desenvolvimentos teóricos e algoritmos adicionais serão necessários para ajudar a iluminá-los como sistemas RC amplamente baseados em dinâmicas intrínsecas de dispositivos podem ser usados em tarefas gerais de aprendizagem de máquina. Esses desenvolvimentos teóricos, juntamente com otimizações contínuas de materiais e dispositivos e avanços de sistemas integrados compostos por arrays memristor diretamente fabricados em circuitos complementares de controle de metal-

óxido-semicondutor e circuitos lógicos⁴⁸, ampliarão ainda mais o apelo dos sistemas RC baseados em memristor para aplicações práticas, como o processamento de dados temporais em tempo real em ambientes com restrição de ^{energia}⁴⁹.

Métodos

Fabricação do dispositivo. O sistema de reservatório à base de memristor foi fabricado em uma estrutura de trave, na qual memristors WOx são formados em cada ponto cruzado. A partir de um substrato com óxido de silício de 100 nm termicamente cultivado em um wafer silicon, 60 nm W foi depositado por sputtering magnetron e padronizado por litografia de feixe eletrônico e gravura de íon reativo usando Ni como uma máscara dura para formar eletrodos de fundo W (BEs) com 500 nm de largura. A máscara de ni hard foi então removida por Gravura molhada. Uma estrutura espaçadora formada ao longo das paredes laterais dos BEs W foi empregada para melhorar o rendimento da fabricação. O espaçador foi padronizado pela deposição de SiO₂ de 250 nm de espessura através de deposição de vapor químico aprimorada pelo plasma, seguido por gravação de íon reativo. A camada de comutação WOX foi então formada nos W BEs expostos por uma rápida ressacimento térmico no ambiente de gás oxigênio a 375 °C por 45 s. Depois, os eletrodos superiores (TEs) com largura de 500 nm foram padronizados por litografia de feixe eletrônico, evaporação do feixe eletrônico de 90 nm Pd e 50 nm Au, e um processo de decolagem. Um processo de gravação de íons reativo foi então usado para remover o WOX entre os TEs para isolar os dispositivos e expor os BEs para contatos elétricos. Finalmente, foram realizados processos de fotolitografia, evaporação do feixe eletrônico e decolagem para formar almofadas de ligação de arame de Au de 150 nm de espessura. Após a fabricação, o chip memristor foi ligado a um porta-chips e montado em uma placa de teste personalizada para testes elétricos. A Fig. 1 suplementar mostra um esquema da estrutura do dispositivo memristor e uma imagem de microscopia eletrônica de varredura ampliada da matriz de 32 × 32 memristor.

Configuração experimental. Para o experimento isolado de reconhecimento de dígito falado, as entradas são convertidas em trens de espigão digitalizados e aplicadas ao reservatório memristor-based. A amplitude/largura dos picos são de 3,0 V/10 μs, e a condutância do dispositivo é lida com pulsos de leitura de 0,6 V/200 μs. O comprimento de um passo de tempo unitário no experimento é de 250 μs.

Para o experimento de previsão de séries temporais, a verdade do solo é obtida pela resolução da equação Mackey-Glass usando o método Runge-Kutta 4 e normalizada na faixa [-0,5, 0,5]. O sinal de entrada $u(t)$ é convertido linearmente em um pulso de programação com 300 μs largura de pulso e amplitude $V \cdot t \cdot 1/4 \cdot u$ $t \cdot \Delta t \cdot 1/8$. A amplitude/largura do pulso de leitura para obter o estado de nó virtual é de 0,6 V/ 200 μs. O comprimento de um passo de tempo unitário é de 150 ms.

Reconhecimento de fala. Formas de onda sonora de dígitos falados isolados (0-9 em inglês) do banco de dados NIST T146 são transformadas em um conjunto de 50 canais de frequência com 40 etapas de tempo usando o modelo de ouvido passivo de Lyon baseado em canais coclear humanos. A saída do modelo de ouvido passivo de Lyon, que é a probabilidade de disparo, é digitalizada com base em se a probabilidade de disparo é maior que 0,5 ou não. Como existem 50 canais de frequência, e n nós virtuais em cada canal, há 50n nós virtuais no total que formam o reservatório. A camada de leitura é, portanto, uma rede de 50n × 10 com 10 saídas representando os 10 dígitos diferentes. Para treinar a rede de leitura, usamos o Kit de ferramentas Python Keras, que fornece uma interface de programação de aplicativos de alto nível para acessar o TensorFlow. Um algoritmo de aprendizagem supervisionado, a regressão softmax, foi usado para treinar a rede de leitura. Uma função softmax é usada como função de ativação da rede de leitura para calcular a probabilidade correspondente às diferentes saídas possíveis. O custo é calculado após uma multientropia categórica. Um método padrão de otimização baseado em gradiente, RMSprop, é usado para minimizar a função de custo e treinar a rede de saída. Após o treinamento da função de leitura utilizando 450 amostras de fala do banco de dados, os dados de teste (das 50 amostras de fala que não estão nos dados de treinamento) são alimentados no reservatório, e a classificação é realizada a partir da rede de leitura com base no estado do reservatório para cada caso de teste. Para evitar que o sistema fosse superequipado para seleções específicas dos dados de treinamento e teste, foi utilizada validação cruzada de 10 vezes. Especificamente, o treinamento e os testes foram repetidos 10 vezes usando 500 amostras de fala (450 amostras para treinamento e 50 amostras para testes), com cada vez tendo uma atribuição diferente de amostras de treinamento e teste. A taxa de reconhecimento foi definida como a média em uma configuração de multivalidação de 10 vezes durante o treinamento. A mesma

função de ativação, função de custo e algoritmo de aprendizagem foram utilizados nos estudos de controle.

Forecasting séries temporais. A série temporal Mackey-Glass é baseada em uma equação diferencial não linear que pode exibir uma ampla gama de comportamentos periódicos e caóticos, dependendo dos valores dos parâmetros. Por exemplo, $\tau < 4,43$ produz um atrator de ponto fixo, $4,43 < \tau < 13,3$ produz um atrator de ciclo de limite estável, $13,3 < \tau < 16,8$ produz um duplo atrator de ciclo de limite e $\tau > 16,8$ produz comportamentos caóticos. Definimos os parâmetros $\beta = 0,2$, $\gamma = 0,1$, $\tau = 18$, $n = 10$. Os dados da série temporal são normalizados na faixa [-0,5, 0,5]. À medida que construímos 20 dispositivos e 50 nós virtuais para cada dispositivo, o estado interno do reservatório é representado por um vetor contendo 1.000 elementos, que é então aplicado à rede de leitura. Um algoritmo de aprendizagem supervisionado, regressão linear, foi usado para treinar a função de leitura. A descida do gradiente estocástico é usada para minimizar o custo, calculado por erro quadrado médio, e treinar a rede de saída. Os resultados dos estados de nó virtual são aplicados à rede de leitura de 1.000×1 para gerar uma saída, o que corresponde ao sinal previsto do próximo passo na série temporal. Na fase de treinamento, 500 etapas do conjunto de treinamento são aplicadas ao reservatório. A saída prevista é comparada com a verdade do solo (no momento passos 51-500), e o erro é calculado e usado para atualizar os pesos na rede de leitura seguindo a regra de aprendizagem de regressão linear.

Disponibilidade de dados

Os dados que suportam as parcelas deste artigo e outros achados deste estudo estão disponíveis no autor correspondente por solicitação razoável.

Recebido: 5 de agosto de 2018; Aceito: 9 de setembro de 2019;

Published online: 14 October 2019

Referências

- Hopfield, J. J. Redes neurais e sistemas físicos com habilidades computacionais coletivas emergentes. *Proc. Natl Acad. Sci. USA* **79**, 2554-2558 (1982).
- Werbos, P. J. Backpropagation through time: o que ele faz e como fazê-lo. *Proc. IEEE* **78**, 1550-1560 (1990).
- Hochreiter, S. & Schmidhuber, J. Longa memória de curto prazo. *Computação Neural*. **9**, 1735-1780 (1997).
- Lukoševičius, M. & Jaeger, H. Reservoir computing se aproxima do treinamento recorrente de rede neural. *Computação. Sci. Rev.* **3**, 127-149 (2009).
- Verstraeten, D., Schrauwen, B. & Stroobandt, D. Técnicas baseadas em reservatórios para reconhecimento de fala. Em *2006 Conferência Conjunta Internacional sobre Redes Neurais (IJCNN)* 1050-1053 (IEEE, 2006).
- Triefenbach, F., Jalalvand, A., Schrauwen, B. & Martens, J.-P. Reconhecimento falso com grandes reservatórios hierárquicos. *Processo Adv. Neural Inf. Syst.* **23**, 2307-2315 (2010).
- Jaeger, H. & Haas, H. Aproveitamento da não linearidade: prever sistemas caóticos e economizar energia na comunicação sem fio. *Ciência* **304**, 78-80 (2004).
- Pathak, J., Hunt, B., Girvan, M., Lu, Z. & Ott, E. Previsão sem modelos de grandes sistemas espaciais caóticos a partir de dados: uma abordagem de computação de reservatório. *Phys. Rev. Lett.* **120**, 024102 (2018).
- Ilies, I. et al. *Stepping Forward Through Echoes of the Past: Forecasting with Echo State Networks* Technical Report http://www.neural-forecasting-competition.com/downloads/NN3/methods/27-NN3_Herbert_Jaeger_report.pdf (Jacobs University Bremen, 2007).
- Sacchi, R., Ozturk, M. C., Principe, J. C., Carneiro, A. A. & Da Silva, I. N. Previsão de entrada de água utilizando a rede estatal eco: um estudo de caso brasileiro. Em *2007 International Joint Conference on Neural Networks (IJCNN)* 2403-2408 (IEEE, 2007).
- Du, C. et al. Computação de reservatório usando memristors dinâmicos para processamento de informações temporais. *Nat. Commun.* **8**, 2204 (2017).
- Sillin, H. O. et al. Um estudo teórico e experimental de redes de interruptores atômicos neuromórficos para computação reservatório. *Nanotecnologia* **24**, 384004 (2013).
- Vandoorne, K. et al. Demonstração experimental de computação de reservatório em um chip fotônico de silício. *Nat. Commun.* **5**, 3541 (2014).
- Torreson, J. et al. Computação neuromórfica com osciladores spintrônicos nanoescala. *Natureza* **547**, 428-431 (2017).
- Chua, L. O. Memristor, o elemento de circuito perdido. *IEEE Trans. Circuit Theory* **18**, 507-519 (1971).
- Waser, R. & Aono, M. Nanoionics resistivamente troca de memórias. *Mater.* **6**, 833-840 (2007).

17. Strukov, D.B., Snider, G.S., Stewart, D. R. & Williams, R. S. O memristor desaparecido encontrado. *Natureza* **453**, 80-83 (2008).
18. Pershin, Y. V. & Di Ventra, M. Neuromórfico, digital e computação quântica com elementos de circuito de memória. *Proc. IEEE* **100**, 2071-2080 (2011).
19. Jo, S.H. et al. Dispositivo de memristor nanoescala como sinapse em sistemas neuromórficos. *O Nano Lett.* **10**, 1297-1301 (2010).
20. Yang, J.J., Strukov, D.B. & Stewart, D. R. Memristive dispositivos para computação. *Nanotecnol.* **8**, 13-24 (2013).
21. Prezioso, M. et al. Treinamento e operação of uma rede neuromórfica integrada baseada em memristors de óxido de metal. *Natureza* **521**, 61-64 (2015).
22. Sheridan, P.M. et al. Codificação esparsa com redes de memristor. *Nanotecnol.* **12**, 784-789 (2017).
23. Apelante, L. et al. Processamento de informações usando um único nó dinâmico como sistema complexo. *Nat. Commun.* **2**, 468 (2011).
24. Maior, L. et al. Processamento de informações fotônicas além de Turing: uma implementação optoeletrônica da computação de reservatórios. *Opt. Express* **20**, 3241-3249 (2012).
25. Brunner, D., Soriano, M.C., Mirasso, C. R. & Fischer, I. Processamento de informações fotônicas paralelas a taxas de dados gigabyte por segundo usando estados transitórios. *Nat. Commun.* **4**, 1364 (2013).
26. Antonik, P., Haelterman, M. & Massar, processador de sinal fotônico inspirado no cérebro de S. Para gerar padrões periódicos e imitar sistemas caóticos. *Phys. Rev. Appl.* **7**, 054014 (2017).
27. Box, G. E. & Pierce, D.A. Distribuição de autocorrelações residuais em modelos de séries de tempo móveis integradas por autoregressiva. *J. Am. Stat. Assoc.* **65**, 1509-1526 (1970).
28. Disse, S.E. & Dickey, D.A. Testando raízes unitárias em modelos médios de ordem desconhecida. *Biometrika* **71**, 599-607 (1984).
29. Kim, K.J. Séries temporais previstas usando máquinas vetoriais de suporte. *Neurocomus* **55**, 307-319 (2003).
30. Kuremoto, T., Kimura, S., Kobayashi, K. & Obayashi, M. Time series forecasting usando uma rede de crenças profundas com máquinas Boltzmann restritas. *Neurocomus* **137**, 47-56 (2014).
31. Zhang, G.P. Uma investigação de redes neurais para previsão linear de séries temporais. *Computação. Oper. Res.* **28**, 1183-1202 (2001).
32. Connor, J.T., Martin, R. D. & Atlas, L.E. Redes neurais recorrentes e previsão robusta de séries temporais. *IEEE Trans. Neural Netw.* **5**, 240-254 (1994).
33. Assaad, M., Boné, R. & Cardot, H. Um novo algoritmo de reforço para uma melhor previsão de séries temporais com redes neurais recorrentes. *Inf. Fusion* **9**, 41-55 (2008).
34. Mirikitani, D.T. & Nikolaev, N. Recursiva bayesiana redes neurais recorrentes para modelagem de séries temporais. *IEEE Trans. Neural Netw.* **21**, 262-274 (2009).
35. Yang, Y. et al. Observação da realização do crescimento de filamentos em memristores resistivos nanoescala. *Nat. Commun.* **3**, 732 (2012).
36. Chang, T., Jo, S. H. & Lu, W. Memória de curto prazo para transição de memória de longo prazo em um memristor nanoescala. *ACS Nano* **5**, 7669-7676 (2011).
37. Wang, Z. et al. Memristors com dinâmica difusiva como emuladores sinápticos para computação neuromórfica. *Mater.* **16**, 101-108 (2017).
38. Texas Instruments-Desenvolvido 46-Word Speaker-Dependent Isolated Word Corpus (TI46) NIST Speech Disc 7-1.1 (Texas Instruments, 1991).
39. Lyon, R.F. Um modelo computacional de filtragem, detecção e compressão na cóclea. *Proc. IEEE-ICASSP'82* **7**, 1282-1285 (1982).
40. Bennett, C. H., Querlioz, D. & Klein, J. O. Spatio-temporal aprendendo com matrizes de nanosynapses analógicas. Em 2017, o IEEE/ACM International Symposium on Nanoscale Architectures (NANOARCH) 125-130 (IEEE, 2017).
41. Frazier, C. & Kockelman, K.M. Chaos theory and transportation systems: instrutivo. *Transp. Res. Rec.* **1897**, 9-17 (2004).
42. Mackey, M.C. & Glass, L. Oscilação e chaos em sistemas de controle fisiológico. *Ciência* **197**, 287-289 (1977).
43. Farmer, J.D. Atratores caóticos de um sistema dinâmico infinitamente dimensional. *Physica D* **4**, 366-393 (1982).
44. Du, C. *Memristors de óxido de metal com dinâmica interna para aplicações neuromórficas* (Universidade de Michigan, 2017).
45. Gallicchio, C., Micheli, A. & Pedrelli, L. Deep reservoir computing: uma análise experimental crítica. *Neurocomus* **268**, 87-99 (2017).
46. Triefenbach, F., Jalalvand, A., Demuyne, K. & Martens, J. P. Acoustic modelagem com reservatórios hierárquicos. *IEEE Trans. Audio Speech Lang. Process.* **21**, 2439-2450 (2013).
47. Gallicchio, C., Micheli, A. & Silvestri, L. Expoentes locais de Lyapunov de redes de eco profundo. *Neurocomus* **298**, 34-45 (2018).
48. Cai, F. et al. Um sistema memristor-CMOS reprogramável totalmente integrado para operações eficientes de multiplicação. *Elétron.* **2**, 290-299 (2019).
49. Tanaka, G. et al. Avanços recentes na computing do reservatório físico: uma revisão. *Netw. neural.* **115**, 100-123 (2019).

Agradecimentos

Reconhecemos discussões inspiradoras com M. Zidan. Este trabalho foi apoiado em parte pela Agência de Projetos de Pesquisa Avançada em Defesa (DARPA) através do prêmio HR001113-2-0015, da Fundação Nacional de Ciência (NSF) através da concessão CCF-1617315, e do Centro de Pesquisa de Aplicações Driving Architectures (ADA), um Centro JUMP copatrocinado da SRC e da DARPA.

contribuições do autor

J.M., W.M. e W.D.L. conceberam o projeto e construíram o quadro de pesquisa. J.M., W.M., F.C. e S.H.L. prepararam as matrizes memristor e construíram o pacote de hardware e software. J.M. e W.M. realizaram as medições de hardware. J.M., W.M., J.H.S. e W.D.L. analisaram os dados experimentais e os resultados da simulação. W.D.L. dirigiu o projeto. Todos os autores discutiram os resultados e implicações e comentaram sobre o manuscrito em todas as etapas.

Interesses concorrentes

Os autores não declaram interesses concorrentes.

informações adicionais

Informações complementares estão disponíveis para o trabalho this na <https://doi.org/10.1038/s41928-019-0313-3>.

Correspondência e pedidos de materiais devem ser endereçados à W.D.L.

As informações sobre reimpressões e permissões estão disponíveis em www.nature.com/reprints.

Nota da editora Springer Nature permanece neutra em relação a reivindicações jurisdicionais em mapas publicados e afiliações institucionais.

© O Autor(s), sob licença exclusiva para Springer Nature Limited 2019