



# Coupled VO<sub>2</sub> Oscillators Circuit as Analog First Layer Filter in Convolutional Neural Networks

Elisabetta Corti<sup>1</sup>, Joaquin Antonio Cornejo Jimenez<sup>1</sup>, Kham M. Niang<sup>2</sup>, John Robertson<sup>2</sup>, Kirsten E. Moselund<sup>1</sup>, Bernd Gotsmann<sup>1</sup>, Adrian M. Ionescu<sup>3</sup> and Siegfried Karg<sup>1\*</sup>

<sup>1</sup> IBM Research Zürich, Rüschlikon, Switzerland, <sup>2</sup> Department of Engineering, University of Cambridge, Cambridge, United Kingdom, <sup>3</sup> Nanoelectronic Devices Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

In this work we present an in-memory computing platform based on coupled VO<sub>2</sub> oscillators fabricated in a crossbar configuration on silicon. Compared to existing platforms, the crossbar configuration promises significant improvements in terms of area density and oscillation frequency. Further, the crossbar devices exhibit low variability and extended reliability, hence, enabling experiments on 4-coupled oscillator. We demonstrate the neuromorphic computing capabilities using the phase relation of the oscillators. As an application, we propose to replace digital filtering operation in a convolutional neural network with oscillating circuits. The concept is tested with a VGG13 architecture on the MNIST dataset, achieving performances of 95% in the recognition task.

**Keywords:** oscillatory neural network, vanadium dioxide, phase-encoding, convolutional neural networks, pattern recognition, relaxation oscillators, coupled oscillators

## OPEN ACCESS

### Edited by:

Peng Li,  
University of California,  
Santa Barbara, United States

### Reviewed by:

Yang Cindy Yi,  
Virginia Tech, United States  
Anup Das,  
Drexel University, United States

### \*Correspondence:

Siegfried Karg  
sfk@zurich.ibm.com

### Specialty section:

This article was submitted to  
Neuromorphic Engineering,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 11 November 2020

**Accepted:** 06 January 2021

**Published:** 11 February 2021

### Citation:

Corti E, Cornejo Jimenez JA, Niang KM, Robertson J, Moselund KE, Gotsmann B, Ionescu AM and Karg S (2021) Coupled VO<sub>2</sub> Oscillators Circuit as Analog First Layer Filter in Convolutional Neural Networks. *Front. Neurosci.* 15:628254. doi: 10.3389/fnins.2021.628254

## INTRODUCTION

Convolutional Neural Networks (CNNs) are the architecture of choice to compute image recognition tasks. Widely used in commercial technology for their recognition accuracy, they are hindered in speed and power efficiency by the frequent access to the memory they need to perform to train a high number of parameters for each convolutional layer in deep networks (Sebastian et al., 2020). The development of neuro-inspired hardware holds the promise of accelerating these algorithms by exploiting in-memory computing concepts and limiting the number of accesses to the memory. A system of coupled oscillator, or Oscillatory Neural Network (ONN) can be used to store and recognize multiple patterns in compact networks. As described in Hoppensteadt and Izhikevich (1999) and Izhikevich (2000), systems of coupled oscillators lock in frequency and establish programmable phase relations that can be used for in-time computing applications. An ONN comprises a system of oscillators, in the role of neurons, connected to each other with synaptic weights, that represent the strength of the oscillators' coupling and the memory of the network. The ONN systems therefore rely on encoding and processing the information with time-delays in the circuits, rather than the amplitude of a signal, therefore being resilient to voltage noise and to scaled power supply.

Exploiting the associative memory capabilities of such networks, tasks as image recognition can be performed. Numerous works have simulated through mathematical and circuit simulations the digit pattern retrieval with different coupled oscillators technologies: (Jackson et al., 2015) shows

20-pixel digit recognition using transition metal oxides and resistive ram technology; (Nikonov et al., 2015; Liyanagedera et al., 2016), perform similar simulations on Spin Torque Oscillators (STOs); (Hölzel and Krischer, 2011) with Van der Pool oscillators. These works are based on storing patterns with  $n \times m$  pixels into an ONN that comprises  $n \times m$  oscillators. To perform the recognition, a distorted pattern of the same pixel size is fed to the ONN, and using the minimum phase attractor of the circuit, the right stored pattern is retrieved. The output is an  $n \times m$  pixels image of the correct digit. Image classifications tasks, however, work quite differently. Taking as an example digit classification through a neural network, an image of  $n \times m$  pixels is fed into the network. The network output is an  $1 \times 10$  array containing the classification probabilities of that image. This classification operation is most commonly performed by convolutional neural networks, that process the image with a series of trainable convolutional filters in the first few layers and achieve recognition after some fully connected layers (Figure 1).

The link between convolutional neural networks and the associative memory capabilities of oscillatory neural networks has so far been explored, to our knowledge, only in Liu and Mukhopadhyay (2018), where an associative memory bank (Hopfield network) replaces the fully connected layers in CNNs. The associative memory is here used to perform a classification with a combination of unsupervised learning and transfer learning techniques. Even though the concept is very interesting and promising, the Hopfield network that this technique envisions comprises between 256 and 2,024 neurons. However, the physical demonstrations of oscillatory neural networks features maximum 100 oscillators as neurons with standard Phase Locked Loop or equivalent CMOS technology (Jackson et al., 2018). The technological challenge in the physical realization of large oscillatory neural networks resides in the complex dynamics of the oscillators' frequency and phase synchronization when the electrical components are affected by variability. This is even more true when the ONNs are built with novel oscillator technology, such as STOs or vanadium-dioxide (VO<sub>2</sub>) oscillators (Romera et al., 2018; Raychowdhury et al., 2019), for which a maximum of 6 oscillators have been connected into a network.

Alternatively, it is suggested that the computing capabilities of small oscillator networks, with up to 10 nodes, can be efficiently exploited to do various image processing tasks, like graph coloring or image saliency processing (Cotter et al., 2014; Tsai et al., 2016). These previous works propose an ONN scheme in which the computation is based on the distance between the input image and the feature to be recognized. For example in Tsai et al. (2016) this distance is encoded in the difference in gate voltages between two transistors which bias a phase transition device. Another popular configuration encodes the distance between the input pattern and the feature to be recognized in a frequency shift between oscillators, which are connected by a fixed coupling. The distance between the two patterns is then calculated on the time the oscillators need to converge to the same frequency (Cotter et al., 2014; Nikonov et al., 2015; Zhang et al., 2019). The implementation of these

concepts does not use the associative memory capabilities of the ONNs to store multiple patterns. Instead, to perform the distance measure, the circuit needs to be reconfigured each time a different feature needs to be recognized. In our work, we propose an implementation of small, fully connected networks, which exploit the associative memory capabilities of an Hopfield network. This allows to store the different features to be recognized in the same network, and enables the recognition of up to 5 different features within one computation performed by the same filter. In addition, in our work we provide the missing link to show how the feature extraction performed by ONNs can be used for image classification tasks. We show the potentiality of the ONN technology for the realization of reconfigurable CNN in hardware, therefore bridging the gap between previous demonstration of ONN pattern retrieval and the industry-standard algorithms.

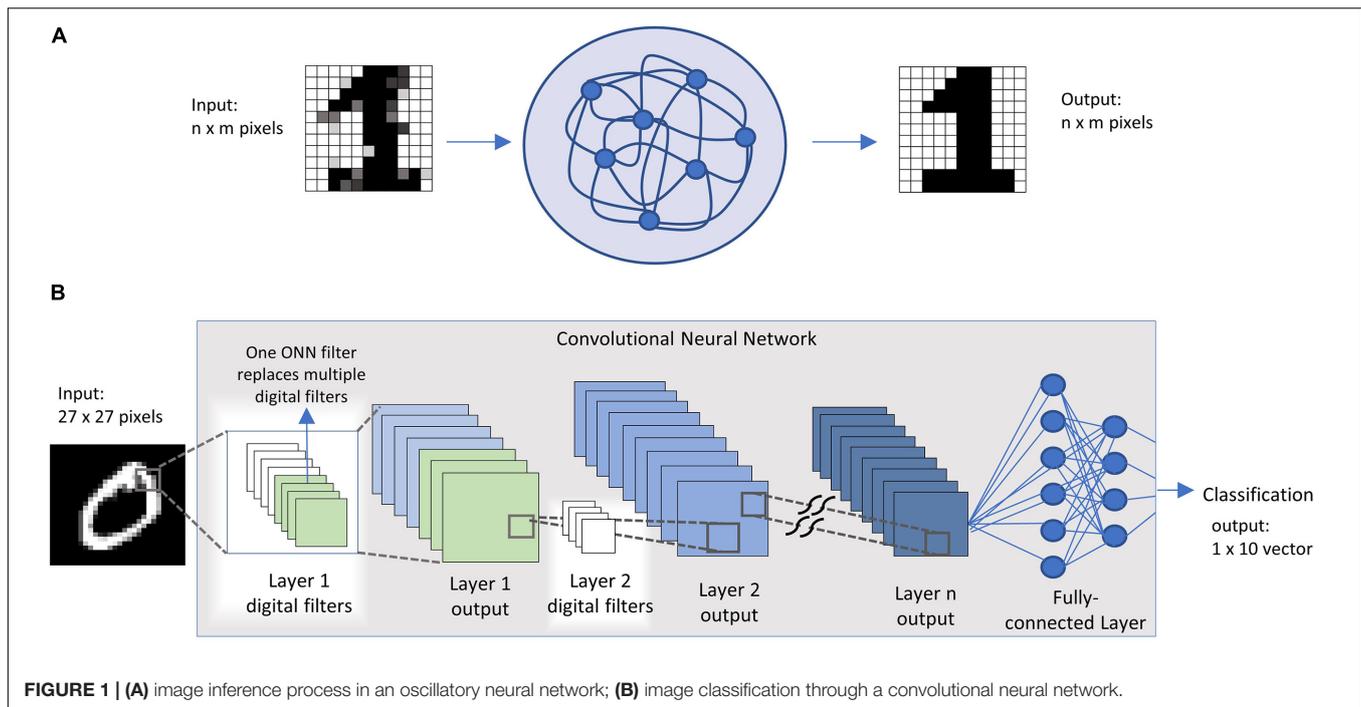
Among the new oscillator technologies, we concentrate our analysis on VO<sub>2</sub> oscillators, as they offer the advantage of realization of very compact oscillators, which can be easily coupled with standard electrical components, allowing for easy reconfigurability of the system (Parihar et al., 2015; Corti et al., 2018). VO<sub>2</sub> based oscillators also offer good scalability perspective and demonstrate operating voltage of less than 1 V and low power consumption ( $\sim 20 \mu\text{W}$  per oscillator) (Shukla et al., 2016).

We exploit the feature extraction capabilities of small networks of VO<sub>2</sub> coupled oscillator to replace digital filters in CNNs (Figure 1B). We fabricate VO<sub>2</sub> oscillators on a Si platform, adopting a crossbar (CB) configuration with scaled device dimensions down to 70 nm. The CB devices exhibit improved variability and reliability compared to co-planar structures and enable the coupling of 4 oscillators. We demonstrate that such a 4-node ONN can memorize and perform 5 different filtering actions of a CNN in a single circuit. Simulations with a  $3 \times 3$  ONN further show how the concept can be applied to replace digital filters in the first layer of a CNN with a VGG-13 inspired architecture and through the adoption of a transfer learning technique. The hybrid CNN-ONN platform has been tested on the MNIST algorithm reaching recognition performances up to 95%. As an outlook, we discuss the benchmark of this technology when extended to all the layers of a CNN, up to the fully connected layers, in comparison with existing hardware and conclude that ONNs might be used as fast and low-power inference machines.

## MATERIALS AND METHODS

### Device Fabrication

VO<sub>2</sub> is a phase change material that presents a volatile, temperature driven insulator to metal transition (IMT). The transition can be triggered by joule heating when a voltage is applied to a VO<sub>2</sub> device (Kim et al., 2010), and it is reversed when the voltage across the device is removed. Given its volatile phase-change characteristics, VO<sub>2</sub> cannot be used as memory element like chalcogenide-based phase change materials (PCM). However, the VO<sub>2</sub> phase transition can be instead



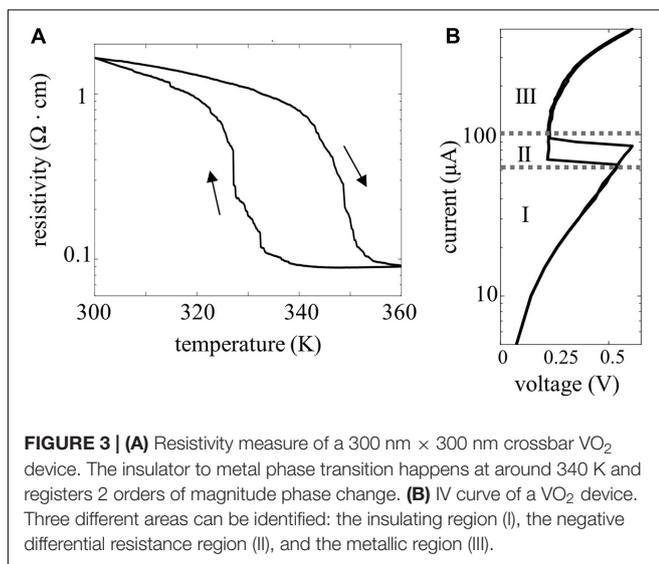
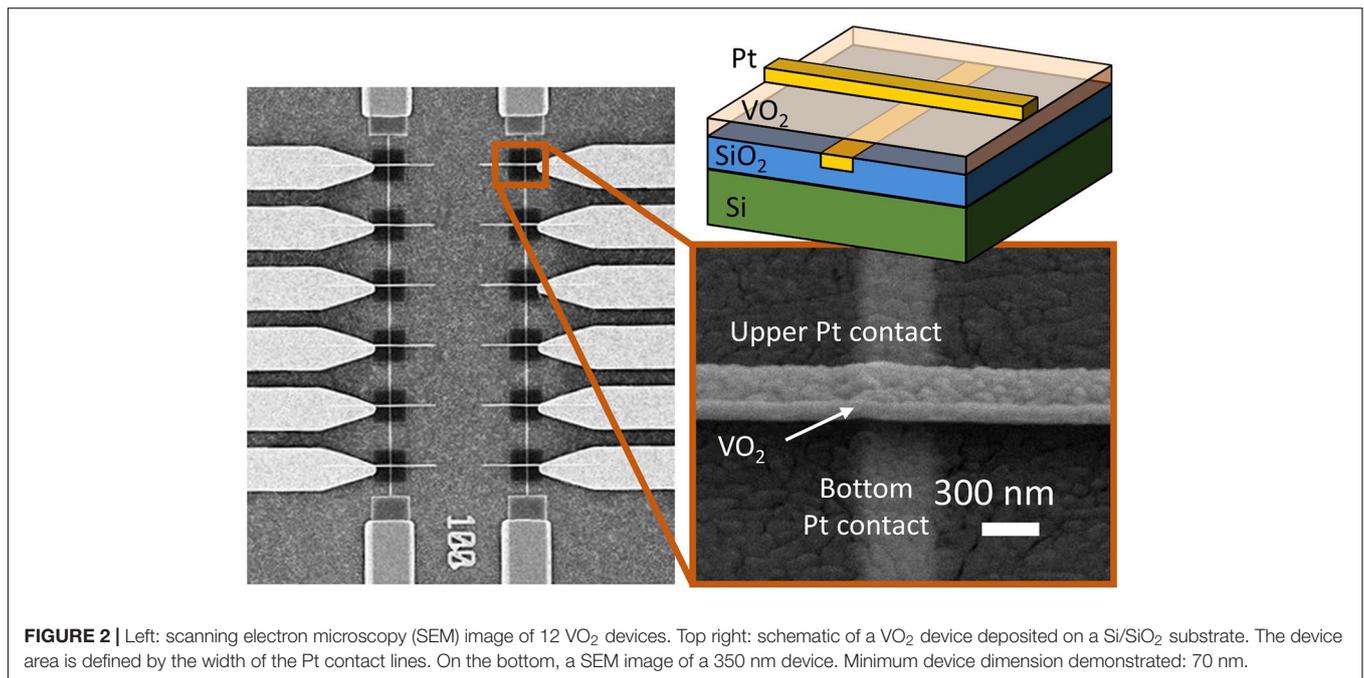
exploited to build very compact oscillators. Other materials have shown similar properties, for example tantalum oxide (Jackson et al., 2015; Sharma et al., 2015) or niobium oxide (Li et al., 2015), however, the near-room temperature phase transition of VO<sub>2</sub> and its proven high endurance up to 10<sup>9</sup> cycles (Shukla et al., 2015) make this material a most favorable choice of oscillators-based technology. VO<sub>2</sub> can be grown crystalline on TiO<sub>2</sub> substrates; however, when deposited on Si, the film forms grains of the average dimension of ~50 nm (Premkumar et al., 2012). In the interest of future integration with electronics we have chosen to focus on integration on silicon in our work.

**Figure 2** shows VO<sub>2</sub> devices fabricated in a CB geometry on a 4<sup>th</sup> Si wafer with a 1 μm thermal SiO<sub>2</sub> layer. Trenches are etched into the SiO<sub>2</sub> substrate and filled with Pt to provide the bottom electrode. Subsequently, an 80 nm thick VO<sub>2</sub> film is grown via atomic layer deposition and post-annealed, resulting in a polycrystalline, granular film (Bai et al., 2020; Niang et al., 2020). Finally, top electrodes are formed using e-beam lithography and Pt evaporation. The smallest device area is 70 nm × 70 nm allowing a very compact design. The resistivity vs. temperature curve (RT) of a 250 nm × 250 nm device is shown in **Figure 3A** and exhibits an insulator-to-metal phase-transition with roughly a two-order of magnitude in resistance change. The step-like RT implies multi-grain transitions, as already shown in previous work (Ruzmetov et al., 2009; Takami et al., 2014; Corti et al., 2019). **Figure 3B** shows the insulator-to-metal and metal-to-insulator transition of an electrically activated device. A current source is used to control the current flowing in the device; a voltmeter is used to measure the voltage at each point. The IV characteristic of this device shows three different operating regions: a first region, in which the device is in its high resistance state, a negative differential resistance regime upon the

phase change, and finally the low resistance region. A crossbar-geometry of VO<sub>2</sub> based-oscillators applications has previously been fabricated with point-probe contacts on TiN on Si, yielding record-speed oscillations performances of 9 MHz, almost an order of magnitude more than what demonstrated with planar structures (Mian et al., 2015). In another more recent work, the oscillating and coupling dynamics of such a vertical structure have been measured and modeled (Tobe et al., 2020). In this work, we further report that the cross-bar geometry yields a better reliability of the devices. In fact, in a previous work we discussed how coplanar devices needed a burn-in cycle to initialize the devices, which sometimes resulted in fatal irreversible changes in morphology (Corti et al., 2019). The crossbar devices do not need a burn-in cycle, improving reliability as virtually all the devices fabricated were able to produce oscillations. Compared to the planar VO<sub>2</sub> structure, the crossbar structure provides improved threshold voltage stability (device-to-device variability lowered from 20% to 10%) and resistance variability (from 10% to ~5%). Compared to other demonstrations on silicon, the improved variability allows for coupling of more oscillator nodes, up to 4. However, to go to larger networks, careful material and device development is necessary to bring this figure down. The devices are tested in temperature-controlled chamber at 320 K and connected in the circuit configurations through external electrical components.

## Oscillatory Neural Network

A single oscillator unit is realized biasing the VO<sub>2</sub> device in the negative differential resistance regime with a series transistor as described in Parihar et al. (2014). When the device is in its insulating state the bias voltage drops mainly across the VO<sub>2</sub>, until the Joule heating brings its local temperature

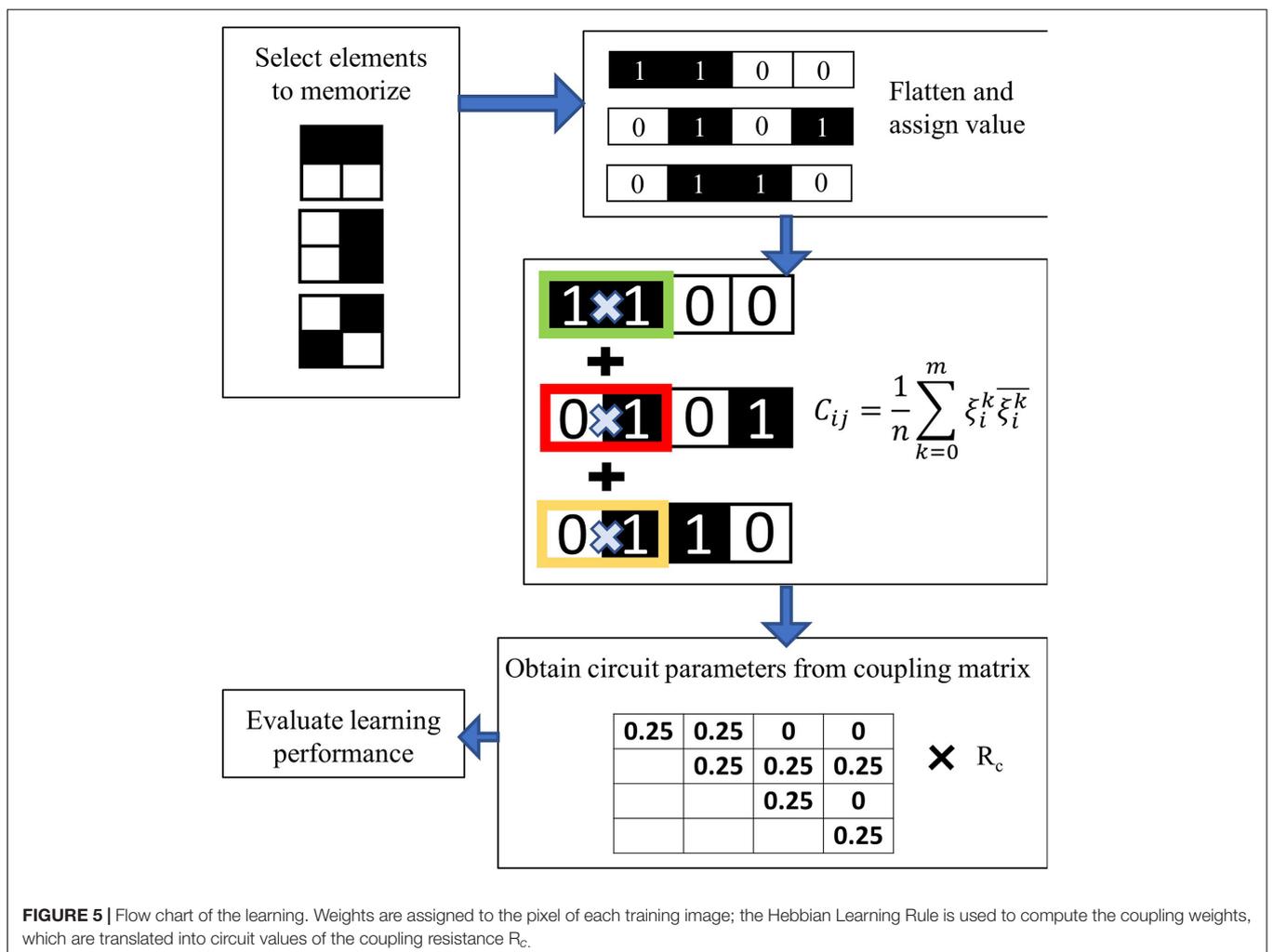
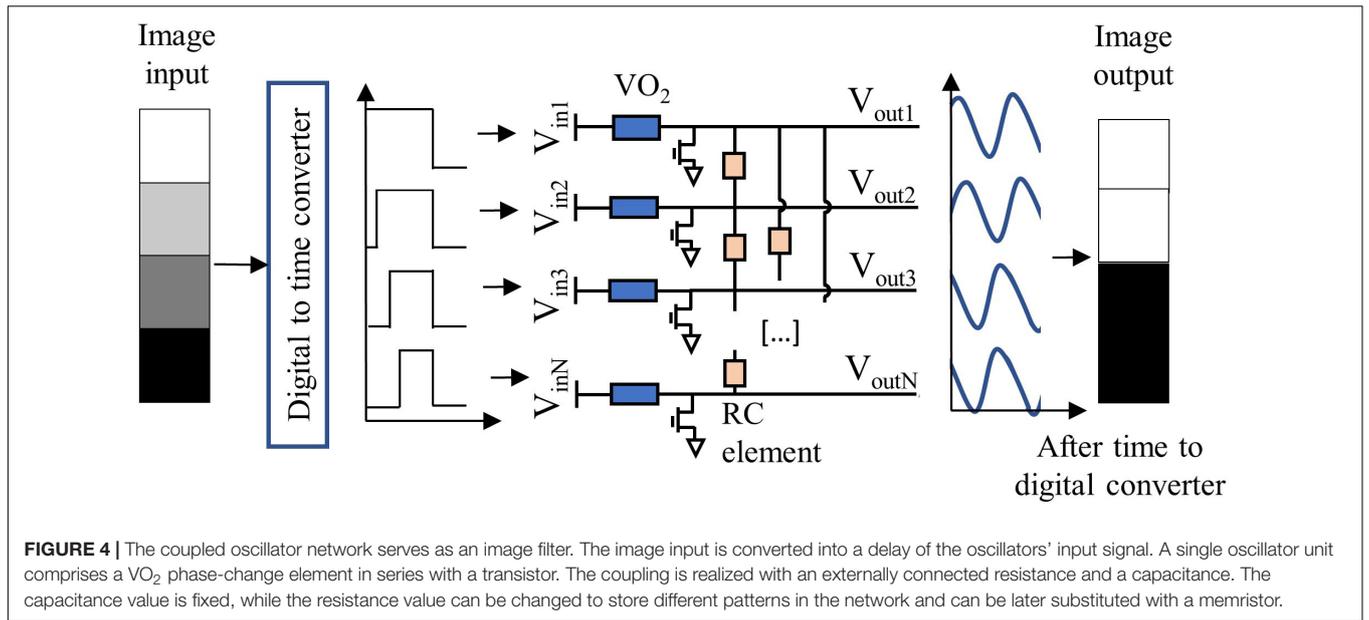


above the phase transition, and the device switches to its metallic state. When the device is in the metallic state, the voltage drops mainly across the series transistor. When the bias is chosen such that the voltage drop does not exceed the upper threshold of the negative-differential resistance regime, Joule heating is reduced. The VO<sub>2</sub> device therefore cools and eventually switches back to its insulating state. The switching between the insulating and metallic state is therefore continuous and self-sustained, originating relaxation oscillations at the drain voltage of the transistor. The oscillators are coupled via resistive and capacitive elements, as shown in **Figure 4**, which ensure frequency and phase-locking of the drain voltage

signals. The strength of the coupling element  $C_{ij}$  that connects oscillator  $i$  with oscillator  $j$  can be calculated starting from the patterns to be memorized, via the Hebbian Learning Rule (HLR) (Hoppensteadt and Izhikevich, 1999):

$$C_{ij} = \frac{1}{n} \sum_{k=0}^m \vartheta_i^k \overline{\vartheta_j^k}$$

where  $n$  is the total number of pixels per each image, or equivalently the number of oscillators in the network,  $\vartheta_i^k$  is the value associated to the pixel  $i$  of pattern  $k$ , and  $m$  is the total number of patterns to be memorized in the ONN (**Figure 5**). These values  $C_{ij}$  are then translated in different values of the coupling resistance  $R_c$  between the oscillators. The memorized patterns appear in the operating ONN as stable phase relations between each oscillator  $i$  and  $j$ . An oscillator in phase with the reference oscillator is translated into a white pixel; an oscillator with 180° phase difference with the reference corresponds to a black pixel. Given  $m$  patterns memorized in the oscillatory neural network, the oscillators can stabilize their phase only according to one of the  $m$  memorized patterns. When the oscillators are initialized to an unstable phase relation, they will relax to the nearest stable ensemble of phase relations, i.e., to the nearest memorized pattern. In this way, from a distorted pattern a memorized pattern is retrieved. In our system in **Figure 4**, the oscillators are initialized to have different phase relations via a delay of the bias voltages to each oscillator. For instance, an oscillator representing a white pixel input is switched on at a time  $t_d = 0$  compared to a reference signal; an oscillator representing a black pixel input is switched on at a time  $t_d = T/2$  compared to the reference signal, with  $T$  indicating the period of one oscillation. Gray-scale values correspond to proportional delays. The output is represented by



the phase of the oscillating transistor drain voltage, compared to a reference. When the network is initialized in this fashion to a phase relation between the oscillators that is different from the stored, stable phase relations, the system relaxes to the nearest stable phase relation, therefore achieving recognition of a pattern. The success of the input time-delay technique for image recognition is explained in detail in Corti et al. (2020). The settling time to the desired output typically varies between 4 and 5 oscillation cycles, nevertheless, after 5 oscillation cycles the phase information becomes stable and can be read-out. The information about the recognized feature is contained solely in the relative phase of the oscillators. In the experiments presented in this paper, the input delay signal is generated through a signal generator unit (National Instruments), the output of the oscillators is acquired by a signal acquisition set-up and the phase calculated with post-processing. The circuit coupling elements are realized with external electrical resistances. As an outlook, the input time-delay can be implemented in hardware via ring oscillators, and the phase-to-digital conversion can be tackled as described in Staszewski et al. (2006). Also, the coupling resistance can be implemented with reconfigurable phase change memories (Boybat et al., 2018).

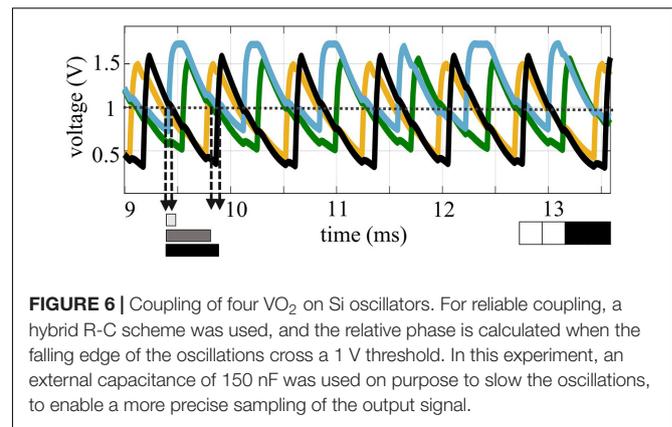
The simulations for the circuit implementation of the ONNs have been done using a Spice simulator. The VO<sub>2</sub> device was simulated with a behavioral model as described in Maffezzoni et al. (2015). TensorFlow™ was used for the CNN and the hybrid ONN-CNN algorithms. The TensorFlow™ code is used to calculate the input delay of the driving voltage of the oscillators, as described above, from an input image, taken from the MNIST dataset. The code then launches the circuit simulation of the ONN, which are conducted in SPICE. The output of the simulation, corresponding to the pattern retrieval computed by the simulated ONN circuit, is then fed back to the TensorFlow™ algorithm as an output image. The image is then processed in the subsequent CNN layers with the TensorFlow™ code.

The choice of the MNIST dataset to perform this computation is justified by the reduced dimensions of the dataset itself. To obtain precise results with the simulation, a high time resolution is required, with very small time-steps for each computation. The simulations of the ONN are therefore rather slow and require an extended simulation time and many computational resources. This problem is not present in the circuit implementation, as the hardware realization is able to perform at frequencies in the order of MHz. We are, however, positive that a similar approach on a more complex dataset will yield the same results here discussed with the MNIST dataset.

## RESULTS

### Four-Coupled Oscillators

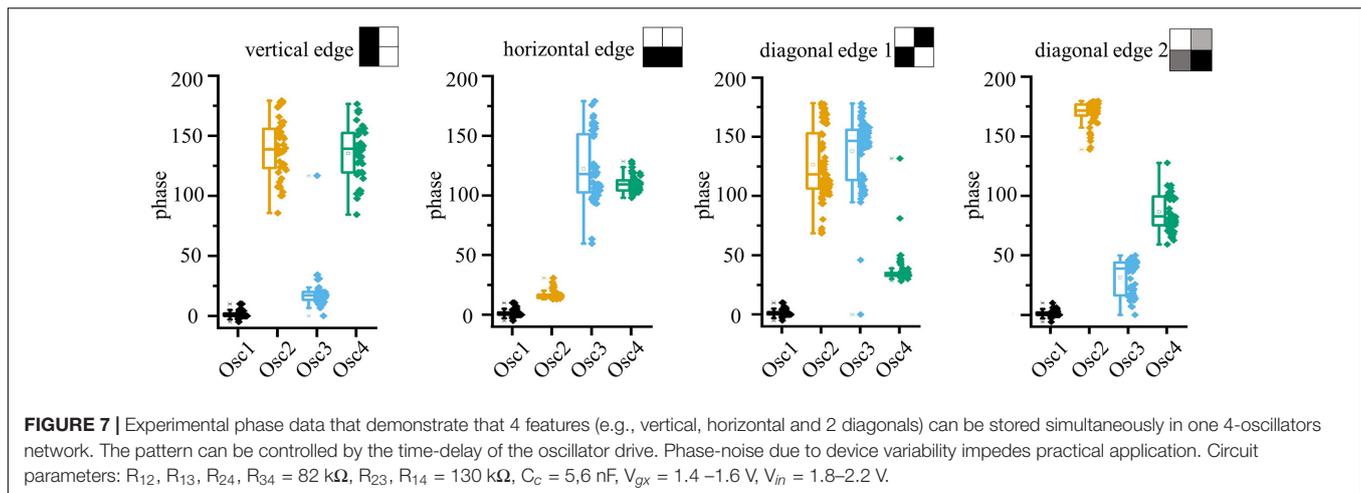
In this section we present a demonstration of four coupled VO<sub>2</sub> oscillators on Si, in which multiple patterns can be memorized. To form relaxation oscillator circuits, the VO<sub>2</sub> resistors on a silicon wafer are coupled through externally connected resistors and capacitances. An example of the measured waveforms of four coupled oscillators is shown in Figure 6. The oscillators



**FIGURE 6 |** Coupling of four VO<sub>2</sub> on Si oscillators. For reliable coupling, a hybrid R-C scheme was used, and the relative phase is calculated when the falling edge of the oscillations cross a 1 V threshold. In this experiment, an external capacitance of 150 nF was used on purpose to slow the oscillations, to enable a more precise sampling of the output signal.

appear to be locked in frequency and the phase relation is calculated taking the distance between the crossing of the 1 V line in the falling edge of the oscillator curves. The coupling network has been programmed to recognize features as in a first layer of a convolutional neural networks. Looking at available analysis of feature extraction in convolutional neural networks (Zeiler and Fergus, 2014), the filters in the first layer commonly select edge features, like borders, diagonal, horizontal and vertical edges. Therefore, for the experimental demonstration, the ONN was trained to store vertical, horizontal and diagonal patterns. The weights of the circuit elements were identified through the Hebbian learning rule. To the best of our knowledge, this is the first demonstration of 4 coupled VO<sub>2</sub> oscillators with memory capabilities realized on a silicon platform. The circuit parameters used for the experiments are:  $R_{12}$ ,  $R_{13}$ ,  $R_{24}$ ,  $R_{34} = 82 \text{ k}\Omega$ ,  $R_{23}$ ,  $R_{14} = 130 \text{ k}\Omega$ ,  $C_c = 5,6 \text{ nF}$ ,  $V_{gx} = 1.4\text{--}1.6 \text{ V}$ ,  $V_{in} = 1.8\text{--}2.2 \text{ V}$ . The different values of gate voltages  $V_g$  and of the input signal  $V_{in}$  are used to achieve similar frequency for the oscillators, and to compensate from intrinsic differences in the devices, which present around 10% of device-to-device variability. The horizontal, vertical and diagonal patterns are identified over multiple experiments, as depicted in Figure 7. In addition, a fourth pattern in which all the oscillators result equally spaced was identified. The measurements are performed assuming Oscillator 1 as the reference oscillator; the phase of the other oscillators is calculated in respect to the crossing of the 1 V threshold of oscillator 1. Therefore, Oscillator 1 has always a phase equal to 0, with a minimal data scattering that is calculated taking into account the variability of the value of the first experimental point that crosses the 1 V line. The other oscillators present a larger scatter, which doesn't impair the clear identification of the various patterns. However, random fluctuation of the oscillations and cross-talk noise hindered the experimental pattern recognition using the input-delay to output phase inference process. This is expected to improve with further process and design optimization of the crossbar devices.

To demonstrate the filtering capabilities of the circuit on an entire image, without suffering from the non-idealities of the experimental demonstration, circuit simulations calibrated



on the experiments were conducted using Spice, reducing the variability of the VO<sub>2</sub> devices from 10% to 5% and therefore increasing the recognition accuracy. In these simulations all the 4 patterns identified in the experiments were also observed; in addition, when the input delays of the circuit were chosen to be all the same, the oscillators in simulations were all oscillating in the in-phase configuration. This is an example of identification of a spurious pattern that was not encoded with the HRL. Spurious patterns arise when the memory capacity of the oscillatory neural network, that is studied to be  $0.15n$  patterns for a  $n$ -oscillator network, is violated (Nishikawa et al., 2004; Follmann et al., 2015). Nevertheless, in such small oscillator networks the spurious patterns can be harvested as additional information. As shown in **Figure 8**, when using the  $2 \times 2$  ONN filter on an image of the MNIST dataset, vertical, horizontal and diagonal edges can be identified. In addition, the background as well as the images parts that have little contrast, can be identified through the in-phase oscillating condition. This demonstrates that a single ONN filter can operate as convolutional feature edge extraction identifying 5 different features. Compared to previous work, the identification of the features does not need to proceed sequentially feature by feature, but it is done in parallel by the same filter. Moreover, the dimensionality of the filter should match the dimensionality of the input, i.e., 2D input arrays such as images are preferably processed using 2D filters. The 4 coupled oscillators system here discussed represents the minimum hardware realization to use ONNs as filters for image feature extraction.

## ONN-CNN

Having shown that our simulations can reproduce experimental behavior, we extend the simulations to explore the use of ONNs in combination with CNNs. The ONN circuit described in section “Oscillatory Neural Network” is simulated with Spice simulations using for the VO<sub>2</sub> device a behavioral model as described in Maffezzoni et al. (2015). The simulations are done with  $3 \times 3$  oscillators ONNs based on parameters

extracted from experimental devices. A convolutional neural network with a structure similar to a VGG-13 is trained on the MNIST dataset with a standard back-propagation algorithm (**Table 1**). The trained weights are used to identify which features are recognized in the first layer of the CNN, that comprises 64 filters with a dimension of  $3 \times 3$ . In our network, as in Zeiler and Fergus (2014), it was also possible to identify multiple filters that selected horizontal, vertical and diagonal edges. We use the Hebbian Learning rule to store the same patterns in a  $3 \times 3$  ONN matrix. The matrix dimension was chosen according to the dimension of the first layer convolution matrixes in the CNN. Ten thousand images from the MNIST dataset have been processed by the ONN matrix with a stride of 2, recognizing in each image vertical, diagonal, horizontal edges and uniform background. As already mentioned, storing of more than  $0.15 n$  patterns, where  $n$  is the number of the oscillators (Follmann et al., 2015), results in the appearance of spurious patterns that can in principle hinder the feature edge extraction process. However, as already discussed for the 4-coupled oscillators experiments, the arising of spurious patterns is not detrimental for feature extraction operations. In the  $3 \times 3$  filter case, we derived the pattern information from 3 key oscillators that oscillate in-phase for each memorized edge. For example, referring to what is depicted in **Figure 9**, each time oscillators 2, 5, and 8 oscillate in-phase a vertical edge is recognized, and similarly for the other edges.

With this technique, a dataset of 10,000 images filtered by the single ONN was calculated, with dimensions  $13 \times 13 \times 5$ , where 5 represent the number of features recognized by the single ONN filter. The dataset was split in 6,000 training images and 4,000 test images.

Subsequently, five filters in the pre-trained CNN that provide the same filtered images were identified and replaced by the ONN with a simple transfer learning process:

1. The 64 CNN filters were convolved with the same images from the MNIST dataset and activated with a Relu function.

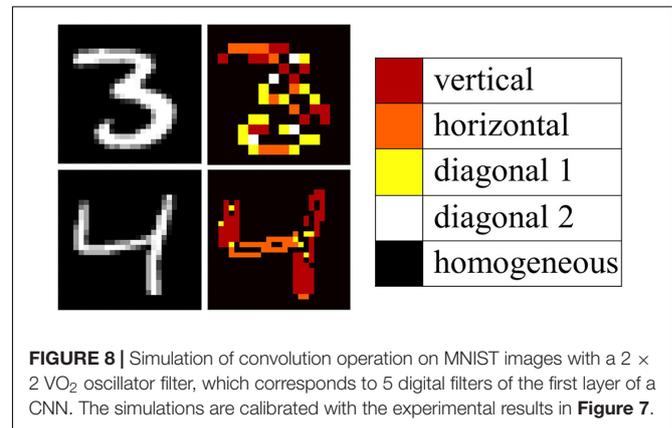
**TABLE 1** | Schematic of the convolutional neural network architecture used in this work for performing the MNIST classification task.

MNIST dataset	27 × 27 × 1,000
ONN-CNN 5 ONN filters + 59 CNN filters	3 × 3 × 64, stride = 2, padding = same
CNN1	3 × 3 × 64, stride = 1, padding = same
Max pool 1	2 × 2, stride = 2, padding = same
CNN 2(×2)	3 × 3 × 128, stride = 1, padding = same
Max pool 2	2 × 2, stride = 2, padding = same
CNN 3(×2)	3 × 3 × 256, stride = 1, padding = same
Max pool 3	2 × 2, stride = 2, padding = same
Fully connected 1	4,096
Fully connected 2	1,000
Fully connected 3	10

The network architecture is inspired by the VGG-13 architecture.

- The CNN-filtered images were compared to the ONN-filtered images calculating the mean square error; the minimum of the mean square error was used to identify the filters from the CNN that can be substituted with the ONN.
- A new dataset is created after the first layer, substituting the images filtered by 5 CNN filters with the 5 filtered images from the ONN.

The remaining neural network layers are trained on the new dataset, achieving a recognition accuracy on the training set of 100% and on the test set of 95%. The original CNN, in comparison, reported better accuracy on the test set, of 97%. The reason for the worsening of the neural network performances is attributed to the cases in which the ONN fails the feature edge extraction. In fact, insufficient training of the ONN (just using HLR) also leads to recognition errors. The implementation of a backpropagation algorithm to the ONN layer would allow to increase the recognition performance in the network. We therefore implemented and tested a backpropagation scheme in our simulations. In **Figure 10** we show an input image feature that should be recognized as a vertical edge. However, when the ONN is trained with the HLR the recognition fails. A cost function  $C = (\varphi_{train} - \varphi_{out})^2/2$  is calculated from on the phases of the desired output  $\varphi_{train}$  and the obtained output  $\varphi_{out}$ . Assuming an exponential dependence of the rising and falling edge of the relaxation oscillator waveforms, the derivative in time can be derived and an improved coupling matrix calculated. During subsequent epochs of this training the phase error is reduced. In the example shown in **Figure 10**, the feature is recognized after 8 epochs of training. While blurred features (allowing 40% gray scale) were only recognized with 30% probability using the untrained ONN, 100% of the features were recognized with the trained ONN. The extension of the backpropagation algorithm to the entire ONN-CNN is yet to be implemented, but is expected to boost the recognition performance. In addition, the direct implementation of the backpropagation algorithm would allow for direct training of a CNN algorithm on an ONN platform and should ultimately result in an increase of the training speed. Despite the reduction in recognition performances, the proposed ONN implementation



**FIGURE 8** | Simulation of convolution operation on MNIST images with a 2 × 2 VO<sub>2</sub> oscillator filter, which corresponds to 5 digital filters of the first layer of a CNN. The simulations are calibrated with the experimental results in **Figure 7**.

allows for a reduction of the number of parameters that need to be trained by the network. In fact, 45 parameters undergo training for 5 CNN filters of 3 × 3 pixels size, however, only 36 parameters need to be trained for a single ONN that performs all filtering actions. The number of parameters to be trained is therefore reduced of 20%; this can represent an important advantage in terms of speed and power consumption when training larger networks. In addition, a further acceleration of the network speed and a further reduction of the number of memory accesses is achieved by the parallel processing of 5 filters from a single ONN unit, whilst in the standard CNN these five convolution actions are performed sequentially.

## Benchmark

In this section we benchmark the convolution operations conducted with the ONN compared to a conventional CPU or GPU. We assume that the first layer of the convolutional neural network presented in this paper is integrally realized via ONN filters operating in parallel. The first layer of the CNN consists of 64 filters of 3 × 3 dimension passing through a 27 × 27 pixel image with a stride of 2, accounting to total of 13 × 13 operations per filter. Assuming that each ONN can perform 5 filtering actions inherently, a total amount of 13 × 13 × 64/5 ≈ 2,200 ONNs is required, which corresponds



**TABLE 2** | Benchmark of the ONN technology against currently available platforms for convolutional neural network applications.

	ONN (current)	ONN (projected)	CPU Intel's Core I9	GPU Tesla V100
Frames/s	$0.6 \times 10^6$	$20 \times 10^6$	$5 \times 10^6$	$600 \times 10^6$
Energy/frame	3.4 $\mu$ J	3 nJ	20 $\mu$ J	500 nJ
TFLOP/s	0.12	4	1	120
TFLOP/s W <sup>-1</sup>	0.06	67	0.01	0.4

roughly to 20,000 oscillator units and 80'000 memristors for implementing the coupling. Assuming a minimum feature size of 100 nm for both the VO<sub>2</sub> oscillators as well as the memristor, the total estimated area would be around 0.001 mm<sup>2</sup>.

For calculating the power consumption of the circuit, we refer to Shukla et al. (2016), Corti et al. (2020), that demonstrate operations of the oscillators at the power  $P = 20 \mu\text{W}$  with a scaled supply voltage  $<1\text{ V}$  and  $f = 3\text{ MHz}$  frequency operation. The total energy for the ONN to process one image with 64 filters at 3 MHz, including the waiting time of 5 oscillating period for the output stabilization, is calculated as

$$P \times f \times 5 = 0.6 \mu\text{J}/\text{frame}$$

Similarly, assuming the mean value of the coupling resistance to be around 100 k $\Omega$ , and the voltage drop across it 0.7 V, the total energy consumption of the memristors is calculated to be 3.4  $\mu\text{J}/\text{frame}$ .

Scaling of the device dimensions, it is envisioned that the VO<sub>2</sub> oscillator could be driven with 1  $\mu\text{W}$  @ 0.3 V at a moderately increased oscillation frequency of 20 MHz. Moreover, through improved processing and resulting device uniformity, the coupling strength could be weakened allowing 1 M $\Omega$  coupling resistance (Shukla et al., 2015). The figure of merit for such a scaled system would improve by 3 orders of magnitude resulting in an energy consumption of 3 nJ/frame.

For conducting the same operation, a standard GPU needs to perform (13  $\times$  13) convolutions  $\times$  64 filters  $\times$  (3  $\times$  3) pixels/filter = 97,344 multiply-accumulation operation, that correspond to around 200,000 flops. In Intel's CPU Core I9, which runs 1 TFLOP/s at 95 W, the total energy accounts for 20  $\mu\text{J}/\text{frame}$ ; in the NVIDIA Tesla V100 GPU, that operates 120 TFLOP/s @ 300 W, the total energy is 500 nJ/frame (Table 2). We can conclude that the ONN system, when built with the current VO<sub>2</sub> technology, is operating now at less power consumption of a conventional CPU, and given the scaling capabilities presented in other works, has the possibility of outperforming the top GPU available on the market. This analysis has been conducted not considering the peripheral circuitry that the ONN system will require, and therefore should be taken just as a projection of the potentiality of this technology and as an indication on the reduction in power consumption that this architecture can bring. Further benchmark should, however, be conducted at a stage when the technology is more advanced, to compare the performances to other

specialized hardware that serve as accelerators for neural networks applications.

## CONCLUSION

A concept for exploiting oscillatory neural networks as hardware accelerators in convolutional neural networks is presented in this paper. A 4-nodes oscillatory neural network was built with scaled VO<sub>2</sub> oscillators' technology on a Si platform. We show that the time-encoded output signal can store up to 5 trained filters and performs the equivalent function of multiple digital convolutional filters in a neural network. We expand the concept to a 3  $\times$  3 VO<sub>2</sub>-ONN trained with Hebbian learning rule and simulate back-propagation for performance optimization. With the 3  $\times$  3 filter and a transfer learning approach, we show that multiple digital filters of a CNN can be trained on a single ONN platform, achieving competitive recognition performances.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

## AUTHOR CONTRIBUTIONS

EC and SK ideated and designed the concepts and experiments proposed in this work. EC was responsible for the device fabrication, the measurements, circuit simulations, and the neural network coding. JC conducted the characterization on the crossbar devices. KN and JR conducted the deposition and annealing of the VO<sub>2</sub> films. KM, BG, and AI were involved in the discussion, experiment design, and editing of the manuscript and provided valuable input at multiple stages of this work. All authors contributed to the article and approved the submitted version.

## FUNDING

We would like to acknowledge the support of our partners and the founding of the HORIZON 2020 NEURONN project (Grant no. 871501) and PHASE-CHANGE SWITCH project (Grant no. 737109).

## REFERENCES

- Bai, G., Niang, K. M., and Robertson, J. (2020). Preparation of atomic layer deposited vanadium dioxide thin films using tetrakis(ethylmethylamino) vanadium as precursor. *J. Vac. Sci. Technol. A* 38:052402. doi: 10.1116/6.0000353
- Boybat, I., Le Gallo, M., Nandakumar, S. R., Moraitis, T., Parnell, T., Tuma, T., et al. (2018). Neuromorphic computing with multi-memristive synapses. *Nat. Commun.* 9:2514. doi: 10.1038/s41467-018-04933-y
- Corti, E., Gotsmann, B., Moselund, K., Ionescu, A. M., Robertson, J., and Karg, S. (2019). Scaled resistively-coupled VO<sub>2</sub> oscillators for neuromorphic computing. *Solid State Electron.* 168:107729. doi: 10.1016/j.sse.2019.10.7729
- Corti, E., Gotsmann, B., Moselund, K., Stolichnov, I., Ionescu, A., and Karg, S. (2018). "Resistive coupled VO<sub>2</sub> oscillators for image recognition," in *Proceedings of the 2018 IEEE International Conference on Rebooting Computing (ICRC)*, McLean, VA (New York, NY: IEEE), 1–7. doi: 10.1109/ICRC.2018.8638626
- Corti, E., Khanna, A., Niang, K., Robertson, J., Moselund, K. E., Gotsmann, B., et al. (2020). Time-delay encoded image recognition in a network of resistively coupled VO on Si oscillators. *IEEE Electron Device Lett.* 41, 629–632. doi: 10.1109/LED.2020.2972006
- Cotter, M. J., Fang, Y., Levitan, S. P., Chiarulli, D. M., and Narayanan, V. (2014). "Computational architectures based on coupled oscillators," in *Proceedings of the IEEE Computer Society Annual Symposium on VLSI, ISVLSI, Tampa, FL* (New York, NY: IEEE Computer Society), 130–135. doi: 10.1109/ISVLSI.2014.87.
- Follmann, R., Macau, E. E. N., Rosa, E., and Piqueira, J. R. C. (2015). Phase oscillatory network and visual pattern recognition. *IEEE Trans. Neural Netw. Learn. Syst.* 26, 1539–1544. doi: 10.1109/TNNLS.2014.23.45572
- Hölzel, R. W., and Krischer, K. (2011). Pattern recognition with simple oscillating circuits. *New J. Phys.* 13:073031. doi: 10.1088/1367-2630/13/7/073031
- Hoppensteadt, F. C., and Izhikevich, E. M. (1999). Oscillatory neurocomputers with dynamic connectivity. *Phys. Rev. Lett.* 82, 2983–2986. doi: 10.1103/PhysRevLett.82.2983
- Izhikevich, E. (2000). *Computing with Oscillators*. Available online at: <http://citeseerx.ist.psu.edu/viewdoc/summary?> (accessed October 19, 2020).
- Jackson, T. C., Sharma, A. A., Bain, J. A., Weldon, J. A., and Pileggi, L. (2015). Oscillatory neural networks based on TMO nano-oscillators and multi-level RRAM cells. 5, 230–241. doi: 10.1109/JETCAS.2015.24.33551
- Jackson, T., Pagliarini, S., and Pileggi, L. (2018). "An oscillatory neural network with programmable resistive synapses in 28 nm CMOS," in *Proceedings of the 2018 IEEE International Conference on Rebooting Computing (ICRC)*, McLean, VA, (New York, NY: IEEE), 1–7. doi: 10.1109/ICRC.2018.86.38600
- Kim, H. T., Kim, B. J., Choi, S., Chae, B. G., Lee, Y. W., Driscoll, T., et al. (2010). Electrical oscillations induced by the metal-insulator transition in VO<sub>2</sub>. *J. Appl. Phys.* 107:023702. doi: 10.1063/1.3275575
- Li, S., Liu, X., Nandi, S. K., Venkatchalam, D. K., and Elliman, R. G. (2015). High-endurance megahertz electrical self-oscillation in Ti/NbOx bilayer structures. *Appl. Phys. Lett.* 106:212902. doi: 10.1063/1.492.1745
- Liu, Q., and Mukhopadhyay, S. (2018). "Unsupervised learning using pretrained CNN and associative memory bank," in *Proceedings of the International Joint Conference on Neural Networks, Rio de Janeiro* (New York, NY: IEEE), 1–8. doi: 10.1109/IJCNN.2018.8489408
- Liyanagedera, C. M., Yogendra, K., Roy, K., and Fan, D. (2016). "Spin torque nano-oscillator based oscillatory neural network" in *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC (New York, NY: IEEE), 1387–1394. doi: 10.1109/IJCNN.2016.77.27360
- Maffezzoni, P., Daniel, L., Shukla, N., Datta, S., and Raychowdhury, A. (2015). Modeling and simulation of vanadium dioxide relaxation oscillators. 62, 2207–2215. doi: 10.1109/TCSI.2015.2452332
- Mian, M. S., Okimura, K., and Sakai, J. (2015). Self-oscillation up to 9 MHz based on voltage triggered switching in VO<sub>2</sub>/TiN point contact junctions. *J. Appl. Phys.* 117:215305. doi: 10.1063/1.4922122
- Niang, K. M., Bai, G., and Robertson, J. (2020). Influence of precursor dose and residence time on the growth rate and uniformity of vanadium dioxide thin films by atomic layer deposition. *J. Vac. Sci. Technol. A* 38:042401. doi: 10.1116/6.0000152
- Nikonov, D. E., Csaba, G., Porod, W., Shibata, T., Voils, D., Hammerstrom, D., et al. (2015). Coupled-oscillator associative memory array operation for pattern recognition. 1, 85–93. doi: 10.1109/JXDC.2015.250.4049
- Parihar, A., Shukla, N., Datta, S., and Raychowdhury, A. (2014). Exploiting synchronization properties of correlated electron devices in a non-boolean computing fabric for template matching. 4, 450–459. doi: 10.1109/JETCAS.2014.2361069
- Parihar, A., Shukla, N., Datta, S., and Raychowdhury, A. (2015). Synchronization of pairwise-coupled, identical, relaxation oscillators based on metal-insulator phase transition devices: a model study. *J. Appl. Phys.* 117, 1–12. doi: 10.1063/1.4906783
- Premkumar, P. A., Toeller, M., Radu, I. P., Adelman, C., Schaeckers, M., Meerschaert, J., et al. (2012). Process study and characterization of VO<sub>2</sub> thin films synthesized by ALD using TEMAV and O<sub>3</sub> precursors. *ECS J. Solid State Sci. Technol.* 1, 169–174. doi: 10.1149/2.009204jss
- Raychowdhury, A., Parihar, A., Smith, G. H., Narayanan, V., Csaba, G., Jerry, M., et al. (2019). Computing with networks of oscillatory dynamical systems. 107, 73–89. doi: 10.1109/JPROC.2018.2878854
- Romera, M., Talatchian, P., Tsunegi, S., Abreu Araujo, F., Cros, V., Bortolotti, P., et al. (2018). Vowel recognition with four coupled spin-torque nano-oscillators. *Nature* 563, 230–234. doi: 10.1038/s41586-018-0632-y
- Ruzmetov, D., Gopalakrishnan, G., Deng, J., Narayanamurti, V., and Ramanathan, S. (2009). Electrical triggering of metal-insulator transition in nanoscale vanadium oxide junctions. *J. Appl. Phys.* 106:083702. doi: 10.1063/1.32.45338
- Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R., and Eleftheriou, E. (2020). Memory devices and applications for in-memory computing. *Nat. Nanotechnol.* 15, 529–544. doi: 10.1038/s41565-020-0655-z
- Sharma, A. A., Li, Y., Skowronski, M., Bain, J. A., and Weldon, J. A. (2015). High-frequency taox-based compact oscillators. *IEEE Trans. Electron Devices* 62, 3857–3862. doi: 10.1109/TED.2015.24.75623
- Shukla, N., Parihar, A., Cotter, M., Barth, M., Li, X., Chandramoorthy, N., et al. (2015). "Pairwise coupled hybrid vanadium dioxide-MOSFET (HVFET) oscillators for non-boolean associative computing," in *Proceedings of the Technical Digest International Electron Devices Meeting IEDM 2015 (February ed.)*, San Francisco, CA (New York, NY: IEEE), 28.7.1–28.7.4. doi: 10.1109/IEDM.2014.7047129
- Shukla, N., Tsai, W. Y., Jerry, M., Barth, M., Narayanan, V., and Datta, S. (2016). "Ultra low power coupled oscillator arrays for computer vision applications," in *Proceedings of the 2016 IEEE Symposium on VLSI Technology: Digest of Technical Papers-Symposium on VLSI Technology*, (Honolulu, HI: Institute of Electrical and Electronics Engineers Inc.). doi: 10.1109/VLSIT.2016.75.73439
- Staszewski, R. B., Vemulapalli, S., Vallur, P., Wallberg, J., and Balsara, P. T. (2006). 1.3 V 20 ps time-to-digital converter for frequency synthesis in 90-nm CMOS. *IEEE Trans. Circ. Syst. II Express Briefs* 53, 220–224. doi: 10.1109/TCSII.2005.858754 \*Cityq
- Takami, H., Kanki, T., and Tanaka, H. (2014). Multistep metal insulator transition in VO<sub>2</sub> nanowires on Al<sub>2</sub>O<sub>3</sub> (0001) substrates. *Appl. Phys. Lett.* 104:23104. doi: 10.1063/1.4861720
- Nishikawa, T., Lai, Y.-C., and Hoppensteadt, F. C. (2004). Capacity of oscillatory associative-memory networks with error-free retrieval. *Phys. Rev. Lett.* 92:108101.
- Tobe, R., Mian, M. S., and Okimura, K. (2020). Coupled oscillations of VO<sub>2</sub>-based layered structures: experiment and simulation approach. *J. Appl. Phys.* 127:195103. doi: 10.1063/5.0001382

- Tsai, W.-Y., Li, X., Jerry, M., Xie, B., Shukla, N., Liu, H., et al. (2016). Enabling new computation paradigms with HyperFET-an emerging device. *IEEE Trans. Multi Scale Comput. Syst.* 2, 30–48. doi: 10.1109/TMSCS.2016.2519022
- Zeiler, M. D., and Fergus, R. (2014). “Visualizing and understanding convolutional networks,” in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, eds D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Cham: Springer Verlag), 818–833. doi: 10.1007/978-3-319-10590-1\_53
- Zhang, T., Haider, M. R., Massoud, Y., and Alexander, J. I. D. (2019). An oscillatory neural network based local processing unit for pattern recognition applications. *Electronics* 8:64. doi: 10.3390/electronics8010064

**Conflict of Interest:** EC, JC, KM, BG, and SK were employed by IBM.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Corti, Cornejo Jimenez, Niang, Robertson, Moselund, Gotsmann, Ionescu and Karg. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.