Multilayer spintronic neural networks with radiofrequency connections

Spintronic nano-synapses and nano-neurons perform neural network operations with high accuracy thanks to their rich, reproducible and controllable magnetization dynamics. These dynamical nanodevices could transform artificial intelligence hardware, provided they implement state-of-the-art deep neural networks. However, there is today no scalable way to connect them in multilayers. Here we show that the flagship nano-components of spintronics, magnetic tunnel junctions, can be connected into multilayer neural networks where they implement both synapses and neurons thanks to their magnetization dynamics, and communicate by processing, transmitting and receiving radiofrequency signals. We build a hardware spintronic neural network composed of nine magnetic tunnel junctions connected in two layers, and show that it natively classifies nonlinearly separable radiofrequency inputs with an accuracy of 97.7%. Using physical simulations, we demonstrate that a large network of nanoscale junctions can achieve state-of-the-art identification of drones from their radiofrequency transmissions, without digitization and consuming only a few milliwatts, which constitutes a gain of several orders of magnitude in power consumption compared to currently used techniques. This study lays the foundation for deep, dynamical, spintronic neural networks.

The recent progress in artificial intelligence (AI) relies on the ability to train deep neural networks that create, layer after layer, more and more meaningful representations of the inputs. In these models, neurons perform a weighted sum of the output of neurons in the previous layer, where each weight represents a synapse, and then apply a nonlinear function to this sum. Neuromorphic computing seeks to reproduce this brain-inspired multilayer architecture and operations on-chip with nanoscale artificial synapses and neurons, in order to place vast amounts of memory at the closest point to computing elements, and hence speed up computing and reduce the overall energy consumption by several orders of magnitude compared to current hardware processors.

Spintronics possesses essential qualities for this purpose, as its flagship component, the magnetic tunnel junction (MTJ), has a high...
We propose and demonstrate an architecture for dense multilayer spinntronics neural networks, exemplified in Fig. 1a. It leverages the intrinsic dynamics and magnetoresistance of MTJs so that these devices emulate both synapses and neurons through successive radiofrequency (RF)-to-d.c. and d.c.-to-RF conversions. A major advantage of this approach is that the resulting neural networks can natively process RF inputs in the wide band of frequencies covered by MTJs (typically 5 MHz to 20 GHz (ref. 28)) and perform, for example, a direct recognition of airborne RF signals, thus avoiding the energy-expensive digitalization step normally used to apply AI techniques to RF signals (ref. 28). We describe the concept of such networks, demonstrate experimentally the operations of its building blocks and show that a hardware network of nine MTJs classifies nonlinearly separable RF inputs with excellent accuracy. We then demonstrate, through physical simulations, the state-of-the-art...
RF spintronic synapses

Synapses are represented by junctions acting as resonators, converting RF to d.c. (refs. 14, 31). A rectified d.c. voltage is produced when an RF signal is injected into a MTJ near its resonance frequency, through magnetoresistance oscillation mixing with the input signal32. The d.c. voltage across the chain of synaptic junctions is plotted versus the d.c. current in neuron 2, for different currents in the field lines (I_{FL}) controlling the weights of the synaptic junctions (I_{J}). I_{J} is set to 16 mA, and I_{FL} varies from dark to light green: I_{FL1} = -15, -17 and -19 mA. The input of neuron 2 is swept while the input of neuron 1 is kept constant at 0 (red) and 7.2 mA (blue). For b and c, the input of neuron 1 is swept while the input of neuron 2 is kept constant at 0 (yellow) and 4.5 mA (brown). For b and c, the vertical dashed lines represent the input d.c. currents below which there is no RF emission. For b and c, the d.c. currents in the field lines are I_{FL1} = 19 mA and I_{FL2} = -18 mA. Neuron 1, neuron 2 and the chain are measured at perpendicular magnetic fields of 270 mT, 690 mT and 193 mT, respectively.

Connecting neurons to synapses

We now demonstrate the basic RF/d.c. interlayer interconnectivity scheme highlighted by yellow dashed lines in Fig. 1a using the neuron and synapse building blocks. We use the two junctions of Fig. 1d as synapses and two other junctions as RF-emitting neurons with matching frequencies f_1 = 268 MHz and f_2 = 383.4 MHz. The synapse junctions are connected in series, receiving the RF outputs of both neurons, amplified and summed with a power combiner (Fig. 2a–c sketches, Methods and Supplementary Fig. 5). We first study the connection between an input neuron and its matching synapse in the chain (Fig. 2a). The bottom panel shows the total voltage of synaptic junctions versus d.c. current I_s at neuron 2’s input, with neuron 1’s input I_{th} set to zero (Supplementary Fig. 4 for neuron 1 response when neuron 2 is off). The measured voltage becomes non-zero above a specific d.c. current, here 3 mA, corresponding to neuron 2’s RF emission threshold. Different curves are obtained by setting various local magnetic fields through the field line above synapse junction 2, indicating that the synaptic junction 2 multiplies neuron 2’s activation function with a tunable weight.

RF spintronic neurons

In a neural network, a neuron performs a weighted sum on concurrently injected RF signals, where each signal frequency is near one of the junctions14,31.

RF spintronic oscillators

The MTJ acts as a neuron that accepts a d.c. input, applies a nonlinear activation function in deep learning: zero below a threshold and proportional to input above it (Supplementary Fig. 3 for neuron characterization). The MTJ as a neuron that accepts a d.c. input, applies a nonlinear activation and broadcasts the computation result as an RF signal. The RF emission frequency from MTJs typically depends on input current due to the anisotropic landscape of magnetization orbits33. We clamp the emission frequency by applying a local magnetic field through a field line (Fig. 1h), experimentally shown in the bottom panel of Fig. 1e (Methods). This frequency clamping ensures that neuron RF oscillations always match the resonant frequency range of corresponding synapses.

RF spintronic synapses

Synapses are represented by junctions acting as resonators, converting RF to d.c. (refs. 14, 31). A rectified d.c. voltage is produced when an RF signal is injected into a MTJ near its resonance frequency, through magnetoresistance oscillation mixing with the input signal32. The d.c. voltage across the chain of synaptic junctions is plotted versus the d.c. current in neuron 2, for different currents in the field lines (I_{FL}) controlling the weights of the synaptic junctions (I_{J}). I_{J} is set to 16 mA, and I_{FL} varies from dark to light green: I_{FL} = -15, -17 and -19 mA. The input of neuron 2 is swept while the input of neuron 1 is kept constant at 0 (red) and 7.2 mA (blue). For b and c, the input of neuron 1 is swept while the input of neuron 2 is kept constant at 0 (yellow) and 4.5 mA (brown). For b and c, the vertical dashed lines represent the input d.c. currents below which there is no RF emission. For b and c, the d.c. currents in the field lines are I_{FL1} = 19 mA and I_{FL2} = -18 mA. Neuron 1, neuron 2 and the chain are measured at perpendicular magnetic fields of 270 mT, 690 mT and 193 mT, respectively.

Connecting neurons to synapses

We now demonstrate the basic RF/d.c. interlayer interconnectivity scheme highlighted by yellow dashed lines in Fig. 1a using the neuron and synapse building blocks. We use the two junctions of Fig. 1d as synapses and two other junctions as RF-emitting neurons with matching frequencies f_1 = 268 MHz and f_2 = 383.4 MHz. The synapse junctions are connected in series, receiving the RF outputs of both neurons, amplified and summed with a power combiner (Fig. 2a–c sketches, Methods and Supplementary Fig. 5). We first study the connection between an input neuron and its matching synapse in the chain (Fig. 2a). The bottom panel shows the total voltage of synaptic junctions versus d.c. current I_s at neuron 2’s input, with neuron 1’s input I_{th} set to zero (Supplementary Fig. 4 for neuron 1 response when neuron 2 is off). The measured voltage becomes non-zero above a specific d.c. current, here 3 mA, corresponding to neuron 2’s RF emission threshold. Different curves are obtained by setting various local magnetic fields through the field line above synapse junction 2, indicating that the synaptic junction 2 multiplies neuron 2’s activation function with a tunable weight.

RF spintronic neurons

In a neural network, a neuron performs a weighted sum on concurrently injected RF signals, where each signal frequency is near one of the junctions14,31.
voltage. Figure 2c shows the same results when the neuron roles are reversed. The rectified voltage increases above ~7 mA, corresponding to neuron 1’s RF emission threshold.

In the Fig. 2a–c plots, dots represent experimental measurements, and lines represent the expected neural network behaviour modelled through

\[ V_{\text{dc}} = a_1 \left( I_{1, \text{dc}} \right) + a_2 \left( I_{2, \text{dc}} \right) + W_1, \]

where \( a_1 \) and \( a_2 \) are neurons 1 and 2’s activation functions, and \( W_1 \) and \( W_2 \) are synaptic weights (model details in Methods). The excellent agreement between experimental data and the neural network model in equation (1) demonstrates that MTJs interconnected through d.c. and RF conversions implement neuron and synaptic operations in a neural network effectively. This result is the experimental demonstration of a functional connection of nanoscale neurons to nanoscale synapses.

### Experimental classification of RF signals

We use this low-noise, high-quality connection scheme to construct the two-layer neural network shown in Fig. 3a, comprising three neurons (two hidden, one output) connected by six synapses. Inputs consist of two RF signals with fixed frequencies \( F_1 \) and \( F_2 \), and varying powers \( P_{1, \text{RF}} \) and \( P_{2, \text{RF}} \). The hardware implementation is in Fig. 3b. We measure d.c. voltages across serially connected synapse junctions and feed neurons with proportional d.c. currents, using external instruments and a computer interface (Methods). To assess the spintronic network performance, we conduct three nonlinear classification tasks illustrated in Fig. 3c’s top panels. The goal is separating dots into two classes in the 2D plane \( (P_{1, \text{RF}}, P_{2, \text{RF}}) \): red if the output neuron emits an RF signal \( (P_{\text{out}, \text{RF}} > 0) \) and blue if not \( (P_{\text{out}, \text{RF}} = 0) \). We train the network offline using the equivalent software network of equation (1), with parameters extracted from the experiment for neurons and synapses. We then set experimental weights close to the trained software network (Methods). The bottom panels in Fig. 3c display corresponding experimental results using offline trained weights and biases (Methods and Supplementary Fig. 6 for data flow example). The black solid lines delineate the target classification. We observe accuracies of 98%, 100% and 95% (97.7% on average), demonstrating the network’s capacity for clean, nonlinear RF signal classification. The small accuracy drop compared to the purely software neural network, which achieves 100% on all tasks, is due to deviations from ideal synaptic operations and imprecisions caused by offline training. Device optimization and chip-in-the-loop training\(^34\) can address these issues (Supplementary Information). However, training requires accessing a wide range of synaptic weights in hardware and accurate device property models. Here this is possible due to spintronic device tunability and the excellent model–experiment fit. The nonlinear solved tasks (that is, red and blue regions not delimited by a single line) demonstrate the hidden neurons’ critical role and interlayer connection.

### Radio-emitter identification by a spintronic neural network

An important advantage of the proposed network architecture is the ability to process different input types. If the first layer is a neuron layer, it takes d.c. inputs, which can connect it to standard electronic circuits. If it’s a synaptic layer, it can directly process RF inputs, as demonstrated

---

**Fig. 3 | Experimental demonstration of the spintronic neural network.**

**a.** Equivalent software network, composed of two RF inputs, two hidden neurons and one output neuron. **b.** Experimental implementation composed of nine MTJs: six of them mimicking synapses, and three neurons. **c.** Nonlinear RF classification task. Each panel shows a 2D plane where each data point has the coordinates \( (P_{1, \text{RF}}, P_{2, \text{RF}}) \) and a colour corresponding to its output class (red or blue).

The black lines are the target boundaries of the classes. The top row shows simulations with the equivalent software network, while the bottom row shows experimental results. Each column is a different task. The three chains and three oscillators are measured at perpendicular fields of 230 mT, 600 mT and 193 mT for the chains, and 270 mT, 690 mT and 415 mT for the oscillators.
with the experimental network. In the latter case, the synaptic layer acts as a spectrum analyser, with extracted results refined and combined in subsequent neuron layers for automatic RF signal classification. Native RF classification has applications in medical imaging 35–37, RF fingerprinting 38 and gesture sensing 39.

We show through simulations that a network of 256 inputs, 128 hidden neurons and 10 outputs can identify drone types directly through their Wi-Fi emissions. Using the drone dataset from the literature 40, we consider nine drone types, drone controllers and a Wi-Fi router (Methods). Figure 4a presents examples of RF signals for different drone types as network inputs. The first layer’s synapses directly process input RF signals and are designed to match input frequencies. By contrast, output layer synapse frequencies match hidden layer neuron frequencies. Each synapse chain processes all output signals from the previous layer. Figure 4b shows classification accuracy evolution versus training epochs for the simulated spintronic network and a pure software network using backpropagation (Methods). We reach maximum accuracies of 99.37 ± 0.48% with the simulated spintronic network and 99.57 ± 0.28% with the standard software network, demonstrating the spintronic neural network classification capabilities in real-world applications, equivalent to standard machine learning methods (Supplementary Information). Chip-in-the-loop training 40 shows that our system can tolerate a device mismatch of up to about 50% of the neuron and synapse resonance width, and about 100% on the neuron activation-function amplitude, before losing 1% accuracy (Supplementary Fig. 7), which is promising for applications.

**Discussion**

The RF spintronic neural network has an important advantage in that it enables applications to process RF inputs natively, without requiring any digitization. Currently, performing the drone classification task described in this work typically requires digitizing the RF inputs with a Universal Software Radio Peripheral, a bulky device consuming about 45 W (ref. 41), as Basak et al. 40 did, and then performing the neural operation on a graphics processing unit that consumes more than a hundred watts 42. By contrast, the spintronic neural network simulated in this work, that natively processes RF inputs, would consume only about 3.4 mW and 0.34 nJ for full processing, orders of magnitude less than conventional technologies (Supplementary Information).

Indeed, with a diameter of 20 nm for each junction, such a large-scale network would consume approximately ten femtojoules per synaptic operation and 100 femtojoules per neural operation 43. This energy consumption takes into account the CMOS RF and d.c. amplifiers needed between each layer of neurons and synapses to maintain signal levels throughout the network (Supplementary Information). It is comparable to or lower than the estimated energy consumption of alternative neuromorphic implementations with memristive or

---

**Fig. 4 | Drone classification by direct processing of their RF emissions.**

- **a**, Spectral representation of the training dataset. Each panel represents the power versus frequency for one class of emitter. Each colour is one example. 
- **b**, Accuracy versus number of training epochs for the simulated spintronic (red) and software (blue) networks. The network architecture is a multilayer perceptron with 256 inputs, 128 hidden neurons and 10 outputs. The uncertainties (shading) correspond to the standard deviation over a hundred trials.
photonics, and several orders of magnitude less than current implementations of software neural networks (Supplementary Information).\(^1\)

Downscaling the MTJs is critical for these future networks to achieve this energy efficiency, as smaller diameters lead to higher frequencies and thus faster computation, as well as to lower threshold currents for neurons and thus lower power consumptions. Furthermore, downscaling combined with the use of amorphous materials will reduce the impact of magnetic grains on the magnetization dynamics and thus improve reliability, decrease the linewidth for better frequency selectivity and bring device behaviour closer to ideal operations\(^{16,40}\). The multifunctionality of spintronic devices will be key to further improve neurons and synapses. Novel effects such as voltage-controlled magnetic anisotropy can be used to increase spin–charge conversion and thus energy efficiency. Non-volatile tuning of the synaptic weights can be achieved by a similar mechanism through modification of the resonance frequencies\(^{12,25,47}\). Alternatively, binary non-volatile synapses exploiting vortex states have been demonstrated\(^{15}\). The emission frequency of the neuronal junctions can be fixed by using junctions with a weak magnetic anisotropy, to avoid the d.c. current-generated magnetic field employed in our proof of concept\(^{10-13}\). Finally, the use of different magnetic fields for each device will be avoided by variability reduction and having devices operating at zero field\(^{10-13}\).

The proposed RF connection scheme based on frequency multiplexing proposed here also constitutes a promising alternative to the crossbar array geometry for densely connecting neurons through synaptic devices with limited resistance variations. Passive crossbar arrays indeed typically require devices with an OFF/ON ratio above 100 to avoid excessive parasitic currents between their columns and rows\(^2\), and they are not adapted to connect low OFF/ON memristors or MTJs\(^3\). By contrast, the presented RF connection scheme does not suffer from sneak paths; by construction, the synaptic chains feeding each output neuron are not interconnected, as the communication between neuron and synapses relies on frequency multiplexing instead of pure wiring, as can be seen in Fig. 1a.

A key element of the frequency multiplexed architecture is having MTJs with different frequencies. While in the present experiment this was achieved by leveraging the variability between devices, this can be engineered by designing nanopillars of different diameters. The frequency range accessible to MTJs and their resonance width sets the maximum number of synaptic junctions in each chain\(^{18}\). With a range of 0.05 to 5 GHz (ref. 28) and a resonance width of 1% of the resonant frequency, we estimate through simulations that a scaled-up network can natively classify RF signals with orders of magnitude less energy than standard approaches, with important applications at stake in telecommunications, medicine, autonomous vehicles and traffic regulation. The intrinsic dynamics of these networks is moreover an exciting opportunity to train them on-chip by algorithms that exploit physical effects for learning\(^{9}\).

**Online content**

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41565-023-01452-w.

**References**

1. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015).

2. Christensen, D. V. et al. 2022 roadmap on neuromorphic computing and engineering. Neuromorphic Comput. Eng. 2, 022501 (2022).

3. Grollier, J. et al. Neuromorphic spinnitronics. Nat. Electron. 3, 360–370 (2020).

4. Worledge, D. C. Spin-Transfer-Torque MRAM: the next revolution in memory. In 2022 IEEE International Memory Workshop (IMW) 1–4 (IEEE, 2022).

5. Safranski, C., Sun, J. Z. & Kent, A. D. A perspective on electrical generation of spin current for magnetic random access memories. Appl. Phys. Lett. 120, 160502 (2022).

6. Lequeux, S. et al. A magnetic synapse: multilevel spin-torque memristor with perpendicular anisotropy. Sci. Rep. 6, 31510 (2016).

7. Kurenkov, A. et al. Artificial neuron and synapse realized in an antiferromagnet/ferromagnet heterostructure using dynamics of spin–orbit torque switching. Adv. Mater. 31, 1900636 (2019).

8. Chen, R. et al. Nanoscale room-temperature multilayer skyrmionic synapse for deep spiking neural networks. Phys. Rev. Appl. 14, 014096 (2020).

9. Song, K. M. et al. Skyrmion-based artificial synapses for neuromorphic computing. Nat. Electron. 3, 148–155 (2020).

10. Mansueto, M. et al. Spintronic memristors for neuromorphic circuits based on the angular variation of tunnel magnetoresistance. Nanoscale 13, 11488–11496 (2021).

11. Zhang, X. et al. Spin-torque memristors based on perpendicular magnetic tunnel junctions for neuromorphic computing. Adv. Sci. 8, 2004645 (2021).

12. Xu, J.-W. et al. A quantum material spintronic resonator. Sci. Rep. 11, 15082 (2021).

13. Kiraly, B., Knol, E. J., van Weerdenburg, W. M. J., Kappen, H. J. & Khajetoorians, A. A. An atomic Boltzmann machine capable of self-adaptation. Nat. Nanotechnol. 16, 414–420 (2021).
14. Leroux, N. et al. Hardware realization of the multiply and accumulate operation on radio-frequency signals with magnetic tunnel junctions. *Neuromorphic Comput. Eng.* [https://doi.org/10.1088/2634-4386/abfca6](https://doi.org/10.1088/2634-4386/abfca6) (2021).
15. Monalisha, P., Kumar, A. P. S., Wang, X. R. & Piramanayagam, S. N. Emulation of synaptic plasticity on a cobalt-based synaptic transistor for neuromorphic computing. *ACS Appl. Mater. Interfaces* **14**, 11864–11872 (2022).
16. Torrejon, J. et al. Neuromorphic computing with nanoscale spintronic oscillators. *Nature* **547**, 428–431 (2017).
17. Prychynenko, D. et al. Magnetic skyrmion as a nonlinear resistive element: a potential building block for reservoir computing. *Phys. Rev. Appl.* **9**, 014034 (2018).
18. Borders, W. A. et al. Integer factorization using stochastic magnetic tunnel junctions. *Nature* **573**, 390–393 (2019).
19. Tsunegi, S. et al. Physical reservoir computing based on spin torque oscillator with forced synchronization. *Appl. Phys. Lett.* **114**, 164101 (2019).
20. Koo, M. et al. Distance computation based on coupled spin-torque oscillators: application to image processing. *Phys. Rev. Appl.* **14**, 034001 (2020).
21. Zahedinejad, M. et al. Two-dimensional mutually synchronized spin Hall nano-oscillator arrays for neuromorphic computing. *Nat. Nanotechnol.* **15**, 47–52 (2020).
22. Ababei, R. V. et al. Neuromorphic computation with a single magnetic domain wall. *Sci. Rep.* **11**, 15587 (2021).
23. Watt, S., Kostylev, M., Ustinov, A. B. & Kalinikos, B. A. Implementing a magnonic reservoir computer model based on time-delay multiplexing. *Phys. Rev. Appl.* **15**, 064060 (2021).
24. Mazza, L. et al. Computing with injection-locked spintronic diodes. *Phys. Rev. Appl.* **17**, 014045 (2022).
25. Zahedinejad, M. et al. Memristive control of mutual spin Hall nano-oscillator synchronization for neuromorphic computing. *Nat. Mater.* **21**, 81–87 (2022).
26. Romera, M. et al. Vowel recognition with four coupled spin-torque nano-oscillators. *Nature* **563**, 230–234 (2018).
27. Romera, M. et al. Binding events through the mutual synchronization of spintronic nano-neurons. *Nat. Commun.* **13**, 883 (2022).
28. Finocchio, G. et al. Perspectives on spintronic diodes. *Appl. Phys. Lett.* **118**, 160502 (2021).
29. Siafarikas, D. & Volakis, J. L. Toward direct RF sampling: implications for digital communications. *IEEE Microw. Mag.* **21**, 43–52 (2020).
30. Farley, B., McGrath, J. & Erdmann, C. An all-programmable 16-nm RFSoC for digital-RF communications. *IEEE Microw. Mag.* **19**, 61–71 (2018).
31. Leroux, N. et al. Radio-frequency multiply-and-accumulate operations with spintronic synapses. *Phys. Rev. Appl.* **15**, 034067 (2021).
32. Tulapurkar, A. A. et al. Spin-torque diode effect in magnetic tunnel junctions. *Nature* **438**, 339–342 (2005).
33. Slavin, A. & Tiberkovich, V. Nonlinear auto-oscillator theory of microwave generation by spin-polarized current. *IEEE Trans. Magn.* **45**, 1875–1918 (2009).
34. Wright, L. G. et al. Deep physical neural networks trained with scaling effect of spin-torque nano-oscillators. *Alp. Adv.* **7**, 056624 (2017).
35. Marković, D., Mizrahi, A., Querlioz, D. & Grollier, J. Physics for neuromorphic computing. *Nat. Rev. Phys.* **2**, 499–510 (2020).
36. Dussaux, A. et al. Large amplitude spin torque vortex oscillations at zero external field using a perpendicular spin polarization. *Appl. Phys. Lett.* **105**, 022404 (2014).
37. Tsunegi, S. et al. High emission power and Q factor in spin torque oscillator consisting of FeB free layer. *Appl. Phys. Express* **7**, 063009 (2014).
38. Choi, J.-G. et al. Voltage-driven gigahertz frequency tuning of spin Hall nano-oscillators. *Nat. Commun.* **13**, 3783 (2022).
39. Martins, L. et al. Non-volatile artificial synapse based on a vortex nano-oscillator. *Sci. Rep.* **11**, 16094 (2021).
40. Jiang, S. et al. Reduced spin torque nano-oscillator linewidth using He” irradiation. *Appl. Phys. Lett.* **116**, 072403 (2020).
41. Divinskii, B., Urazhdin, S., Demokritov, S. O. & Demidov, V. E. Controlled nonlinear magnetic damping in spin-Hall nano-devices. *Nat. Commun.* **10**, 5211 (2019).
42. Sethi, P. et al. Compensation of anisotropy in spin-Hall devices for neuromorphic applications. *Phys. Rev. Appl.* **19**, 064018 (2023).
43. Jenkins, A. S., Alvarez, L. S. E., Freitas, P. P. & Ferreira, R. Digital and analogue modulation and demodulation scheme using vortex-based spin torque nano-oscillators. *Sci. Rep.* **10**, 11181 (2020).
44. Jung, S. et al. A crossbar array of magnetoresistive memory devices for in-memory computing. *Nature* **601**, 211–216 (2022).
45. Craven, M. P., Curtis, K. M. & Hayes-Gill, B. R. Frequency division multiplexing in analogue neural network. *Electron. Lett.* **27**, 918–920 (1991).
46. Leroux, N. et al. Convolutional neural networks with radio-frequency spintronic nano-devices. *Neuromorphic Comput. Eng.* **2**, 034002 (2022).
47. Noh, S. et al. Spin dynamics in ferromagnetic resonance for nano-sized magnetic dot arrays: metrology and insight into magnetization dynamics. *IEEE Trans. Magn.* **47**, 2387–2390 (2011).
48. Litvinenko, A. et al. Ultrafast GHz-range swept-tuned spectrum analyzer with 20 ns temporal resolution based on a spin-torque nano-oscillator with a uniformly magnetized ‘free’ layer. *Nano Lett.* **22**, 1874–1879 (2022).
49. Kurokawa, Y. et al. Ultra-wide-band millimeter-wave generator using spin torque oscillator with strong interlayer exchange couplings. *Sci. Rep.* **12**, 10849 (2022).
50. Bonetti, S., Muduli, P., Mancoff, F. & Åkerman, J. Spin torque oscillator frequency versus magnetic field angle: the prospect of operation beyond 65 GHz. *Appl. Phys. Lett.* **94**, 102507 (2009).
51. Sethi, P. et al. Compensation of anisotropy in spin-Hall devices for neuromorphic applications. Nature Nanotechnology | Volume 18 | November 2023 | 1273–1280.
52. Jenkins, A. S., Alvarez, L. S. E., Freitas, P. P. & Ferreira, R. Digital and analogue modulation and demodulation scheme using vortex-based spin torque nano-oscillators. *Sci. Rep.* **10**, 11181 (2020).
53. Jung, S. et al. A crossbar array of magnetoresistive memory devices for in-memory computing. *Nature* **601**, 211–216 (2022).
54. Craven, M. P., Curtis, K. M. & Hayes-Gill, B. R. Frequency division multiplexing in analogue neural network. *Electron. Lett.* **27**, 918–920 (1991).
55. Leroux, N. et al. Convolutional neural networks with radio-frequency spintronic nano-devices. *Neuromorphic Comput. Eng.* **2**, 034002 (2022).
56. Noh, S. et al. Spin dynamics in ferromagnetic resonance for nano-sized magnetic dot arrays: metrology and insight into magnetization dynamics. *IEEE Trans. Magn.* **47**, 2387–2390 (2011).
57. Litvinenko, A. et al. Ultrafast GHz-range swept-tuned spectrum analyzer with 20 ns temporal resolution based on a spin-torque nano-oscillator with a uniformly magnetized ‘free’ layer. *Nano Lett.* **22**, 1874–1879 (2022).
58. Kurokawa, Y. et al. Ultra-wide-band millimeter-wave generator using spin torque oscillator with strong interlayer exchange couplings. *Sci. Rep.* **12**, 10849 (2022).
59. Bonetti, S., Muduli, P., Mancoff, F. & Åkerman, J. Spin torque oscillator frequency versus magnetic field angle: the prospect of operation beyond 65 GHz. *Appl. Phys. Lett.* **94**, 102507 (2009).
60. Khymyn, R. et al. Ultra-fast artificial neuron: generation of picosecond-duration spikes in a current-driven antiferromagnetic auto-oscillator. *Sci. Rep.* **8**, 15727 (2018).
61. Chakravarty, A. et al. Supervised learning of an opto-magnetic neural network with ultrashort laser pulses. Appl. Phys. Lett. 114, 192407 (2019).

62. Scellier, B. & Bengio, Y. Equilibrium propagation: bridging the gap between energy-based models and backpropagation. Front. Comput. Neurosci. 11, 24 (2017).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

MTJ sample preparation

MTJ films have a stacking structure as follows: buffer (5 Ta / 50 CuN / 5 Ta / 50 CuN / 5 Ta / 5 Ru) / antiferromagnet (6 IrMn) / synthetic antiferromagnet (2.0 CoFeO/0.7 Ru / 2.6 CoFeO/B2O) / tunnel barrier (-1 MgO (resistance x area product in the 3–12 Ω μm2 range)) / free layer (2.0 CoFeO/B2O / 0.2 Ta / 7 NiFe) / 10 Ta / 7 Ru, with all thicknesses in nanometres.

The MTJ stack is deposited by magnetron sputtering using the Timaris MTM tool. After electron-beam lithography with a negative AR-N 7520.18 resist, the pillars are defined by ion beam milling. The free layer is a composite of CoFeB and NiFe. The CoFeB is amorphous as deposited and crystalizes into body-centred cubic 100 during annealing (330 °C, 1 tesla, 2 hours), which is necessary to ensure high-quality MgO growth, and the NiFe is a soft magnetic material that dominates the dynamic behaviour. The 0.2 nm Ta is deposited to stop the NiFe face-centred cubic 111 from influencing the crystalization at the CoFeB. The pillars have a nominal diameter of 350 nm.

The free layer has a magnetic vortex as the ground state. In a small region called the vortex core, the magnetization spirals out of plane. Under d.c. or RF current injection and the action of the Oersted field and spin transfer torques, the core of the vortex gyrates at a frequency in the range of 150 MHz to 450 MHz for the oscillators and resonators we used here.

The typical resistance versus in-plane field profile is presented in Supplementary Fig. 1. The tunnel magnetoresistance is 72%.

Field line for local magnetic field application

The field line is a 300-nm-thick and 3-μm-wide AlSiCu stripline, deposited 600 nm above the free layer of the MTJ. The local magnetic field generated by d.c. current injection in this stripline is collinear with the free layer and about 0.05–0.1 mT mA−1.

Characterization of the synaptic and neuron devices of the fully spintronic neural network

Synaptic chains. In the network, we used three chains of two serially connected synaptic junctions each. The chains, presented in Supplementary Fig. 2, are measured under out-of-plane magnetic fields of 193 mT, 230 mT and 600 mT.

Neurons. In the network, we used three neurons: neurons 1, 2 (hidden) and 3 (output), which are fully characterized in Supplementary Fig. 3. The measurements are performed under out-of-plane magnetic fields of 270 mT, 690 mT and 415 mT, respectively. The frequencies of neurons 1 and 2 are fixed at 266.46 ± 1.36 MHz, 384.09 ± 0.90 MHz and 254.68 ± 0.19 MHz, respectively. The value of the d.c. current to inject into the field line (IFL) at different values of d.c. current into the device (Pd) is found through prior characterization and interpolation. These measurements are done with no amplification for neurons 1 and 2 and with 11 dB amplification for neuron 3. The d.c. current input is capped in order to prevent the breakdown of the tunnel barrier. This restriction is taken into account in the spintronic neural network simulations.

Interconnection between two neurons and two synapses. Neuron 1, neuron 2 and the chain of synapses are measured in the same conditions as in their characterization, described in the previous section. We use a dedicated experimental set-up that enables us to apply different magnetic fields to different devices simultaneously, using pairs of permanent magnets. The RF signals are propagated from the neurons to the synapses through electrical wires. Each of the parts (the two neurons and the chain) is connected to SMA connectors through wire-bonding. Then each neuron is connected to a circulator (to protect the device from reflections), and then to an RF amplifier. The amplifiers are of 6 dB and 11 dB for neurons 1 and 2, respectively. The amplifications were chosen such that the output powers of the two hidden layer neurons were comparable in the final network (maximum powers of 3.4 μW and 3.9 μW, respectively). The two resulting signals were combined by a power combiner (6 dB loss) and sent to the chain. The d.c. inputs are provided via a multi-channel source measurement unit, to the neurons and their field lines, as well as to field lines of the synapses in order to vary their weights.

In total, 49 sets of weights were obtained (Ih1 from 16 to 19 mA with 0.5 mA steps, and Ih2 from 15 to 18 mA with 0.5 mA steps). For each set of weights, the experimental data are compared with the following model:

\[ V_{dc} = V_1(I_{dc}^{FL1}) + V_2(I_{dc}^{FL2}) \]

with

\[ V_i(I_{dc}) = \begin{cases} 
0 & \text{if } I_{dc} < I_{th,i} \\
W_i \times I_{dc}^{FL} + c_i & \text{if } I_{dc} \geq I_{th,i} 
\end{cases} \]

where the weights \( W_i \), the constants \( c_i \) and the threshold currents \( I_{th,i} \) are constants extracted from the experimental data. The \( I_{dc}^{FL} \) values are the input d.c. currents to each neuron.

For the combination of weights shown in Fig. 2b,c (Ih1 = 19 mA and Ih2 = 18 mA), the root-mean-square error over all oscillator currents is 0.17 μV, that is, 2.3% of the range. For all 49 measured combinations of weights, the root-mean-square error is 0.30 μV, that is, 3.6% of the range.

Two-layer fully spintronic hardware neural network

The three synaptic chains and three neurons are measured in the conditions of their characterization described in the previous sections. The inference with the fully spintronic neural network is conducted as follows, and as described in Supplementary Fig. 5:

- Each input is a pair of two RF powers (\( P_{1}^{RF}, P_{2}^{RF} \)). The input powers each range from 0.5 to 3 μW. Their respective frequencies are fixed at \( f_{1} = 220 \) MHz and \( f_{2} = 400 \) MHz. The input RF signals are generated by a computer-controlled RF source.
- The two input RF powers are combined through a power combiner, and the resulting signal is injected in the two synaptic chains of the first layer through a power splitter.
- The two output d.c. voltages of the synaptic chains are recorded by the computer. The computer converts each d.c. voltage into a d.c. current value (that is, multiplication by a scaling factor). A bias is added to each d.c. current value. The computer controls the resulting d.c. current applied by current sources to each neuron.
- The RF outputs of the two neurons are amplified, combined and injected into the synaptic chain of the second layer (as described in the ‘Interconnection’ section in the Methods).
- The output d.c. voltage of the synaptic chain is recorded by the computer and converted into a d.c. current (scaling and bias), and injected into the output neuron by a d.c. current source.
- The output RF power emitted by the output neuron is amplified by 11 dB and recorded on the computer through a spectrum analyser.
- The output of the network is the presence or absence of RF emission by the output neuron (Fig. 3c).

Experimental RF signal classification

The equivalent software neural network is composed of the analytical models following equation (1) of the synaptic chains (weighted sum) and the neuron activation functions, as well as one bias for each neuron.
We define three different nonlinear tasks. For each task the procedure is as follows. We define boundaries in the (power 1, power 2) two-dimensional plane of input power values, marked by black lines, that delimit two regions (class 0 and class 1). We train the equivalent software neural network using the neural network library PyTorch through backpropagation (using Adam as the optimizer and mean-square error as the loss), so that the output of the network is 1 in one region (red) and 0 in the other (blue). Thus, the inputs are (power 1, power 2) couples and the targets are 0 or 1. The training provides a set of ideal weights for the synapses. We set the d.c. currents in the field lines of the synaptic chains so that the synaptic weights match the ideal weights as closely as possible. We then perform the experimental classification tasks by injecting (power 1, power 2) input pairs into the hardware network.

Drone classification dataset
The drone classification dataset is composed of ten classes of signals. The first nine classes consist of signals collected from drone radio controllers, and the last one, from Wi-Fi signals. The signals, originally all in the Wi-Fi band, were recorded using a Universal Software Radio Peripheral with down-conversion to the 0–100 MHz range, and made public by Basak et al.40. The signals collected by Basak et al. are spectrograms of size 256 by 256. We have averaged these spectrograms over the fast Fourier transform frame numbers, representing the time, and obtained discrete Fourier transforms with an averaged amplitude. We have scaled this amplitude over the whole dataset from 0 to 1 uW. For better matching with the physical network, we translate the signals to the 20–120 MHz range. The resulting dataset from 0 to 1 uW. For better matching with the physical network, we translate the signals to the 20–120 MHz range. The resulting dataset, used in this Article, is available online: https://doi.org/10.5281/zenodo.7646236.

Physical models used for the simulated neural network

Weighted sum. In our case, the input spectrum is composed of a finite number $N_{\text{input}}$ of equidistant input frequency bins so that the total d.c. voltage $V$ rectified by one resonator of resonance frequency $f_i$ is the sum of the individual effects of each frequency:

$$V(f_i) = \sum_{i} P(f_i) G(f_i, f_{\text{res}})$$

where $P(f_i)$ is the input power and $G(f_i, f_{\text{res}})$ is the rectification coefficient:

$$G(f_i, f_{\text{res}}) = \frac{2a f_{\text{res}}^2 (f_i - f_{\text{res}}) k_{\text{SD}}}{(a f_{\text{res}}^2)^2 + (f_i - f_{\text{res}})^2}$$

where $a = 0.01$ is the Gilbert damping and $k_{\text{SD}}$ is the spin-diode sensitivity.

We consider $N_{\text{rec}}$ resonators per synaptic chain, where each resonator $k$ of a chain $j$ will rectify the input signal. As a result, the rectification function can be written conveniently using a three tensor:

$$G_{jk} = \frac{2a f_{\text{res}}^2 (f_i - f_{\text{res}}) k_{\text{SD}}}{(a f_{\text{res}}^2)^2 + (f_i - f_{\text{res}})^2}$$

The weight applied by the synaptic chain $j$ to the input $P(f_i)$ corresponds to the total rectification done by all the resonators of the chain $j$ for the frequency component $f_i$:

$$W_{ij} = \sum_{k} (-1)^k G_{jk}$$

where $(-1)^k$ comes from the fact that the resonators are electrically connected in a ‘head-to-tail’ manner. Indeed, when resonators are connected ‘head-to-tail’ (as shown in Fig. 1g), their low-frequency finite components add up and create a voltage drift (as observed in Fig. 1d) that increases with the number of resonators. For chains with a large number of resonators, it is therefore preferable to connect the resonators ‘head-to-head’, as depicted in Fig. 1a.

The output voltages are expressed as a classical neural network weighted sum:

$$V_j = \sum_{i} P_i W_{ij} + b_j$$

We conveniently write this equation using vectors and matrices:

$$V = PW + b.$$  (3)

The voltage bias $b$ is actually composed of two terms:

$$b = V_{\text{chains}} + V_{\text{layer}}$$

where $V_{\text{chains}}$ is the vector containing the voltage biases of the chains (learning parameters) and $V_{\text{layer}}$ is a vector of the same dimension whose components correspond to an additional constant voltage bias (hyperparameter) used to improve the performance of the network. This constitutes the first operation of the network, that is, the weighted sum.

For the numerical values, $N$ is fixed by the dataset to 256, $a = 0.01$ and $k_{\text{SD}} = 8.8 \times 10^3 \mu V \mu W^{-1}$.

Activation function. The activation function is modelled by the response of a spin-torque nano-oscillator to a d.c. current $I_{\text{dc}}$. If this current is superior to a threshold current $I_{\text{th}}$, the oscillator enters into a stationary precession regime, and its normalized magnetization oscillation power can be modelled by the following:

$$P_j = A |c|^2 \left( \frac{R^2}{R_s} \right)^2 R (p_{\text{c}})^2,$$

where $A$ is a scaling parameter, $R$ is the resistance, $\frac{k_{\text{SD}}}{R}$ is the tunnel magnetoresistance and $\beta_s$ is the shape factor. We simplify this expression by considering $\frac{k_{\text{SD}}}{R} \beta_s \approx 1$:

$$P_j = A |c|^2 R (p_{\text{c}})^2.$$

Additionally, the input current of each oscillator of the hidden layer is clamped, $I_{\text{dc}} = 4 I_{\text{th}}$, to mimic the security preventing the oscillators from breaking. The resulting output powers of the hidden layer are therefore

$$p_{\text{hidden}} = \left\{ \begin{array}{ll} A |c|^2 I_{\text{dc}} \beta_s^2 & \text{if } p_{\text{c}} \geq I_{\text{th}} \\ A |c|^2 R (p_{\text{c}})^2 \beta_s^2 & \text{if } I_{\text{th}} > p_{\text{c}} > I_{\text{th}} \\ 0 & \text{if } p_{\text{c}} \leq I_{\text{th}} \end{array} \right.$$
To obtain the vector of currents $I^{dc}$ from the output voltages obtained with equation (3), we use a voltage-to-current ratio $g_m$ mimicking a transconductance amplifier:

$$I^{dc} = V_g.$$

These previous equations form the building blocks of the inference of the physical deep neural network.

**Neural network architecture.** We consider the same architectures for the equivalent software and simulated physical neural networks: a two-layer neural network with 256 inputs (corresponding to the 256 frequency bins from 20 to 120 MHz), 128 hidden neurons and 10 outputs (corresponding to the 10 classes).

While for the equivalent software neural network, the trained parameters are the weights and biases, for the simulated physical neural network, the trained parameters are the resonance frequencies of all the resonators $f_{res}^{jk}$ and the output voltage biases $V_{chains}^{j}$.

**Model selection.** For the model selection, we used a stratified fivefold cross-validation method. For training we used the cross-entropy loss and the Adam optimizer.

The hyperparameters of each final model were selected with the use of a hyperparameter optimization procedure based on the Optuna framework. For the software model, there is only one hyperparameter, which is the learning rate. For the physical model, the hyperparameters are the learning rate as well as the additive bias $V^{layer}$ on the output voltages of the first synaptic layer and the voltage-to-current ratio $g_m$ converting the output voltages of the first synaptic layer into currents.

The number of trials (that is, the set of hyperparameters tested) was set to 100. The number of epochs was set to 100, and the training/validation procedure was repeated 10 times in order to use the resulting mean accuracy as an objective function for the Optuna optimizer.

For the software model, the resulting learning rate was $\eta = 0.06$. For the physical model, the resulting learning rate was $\eta = 1.07 \times 10^{-5}$, the voltage-to-current ratio $g_m = 1.81 \mu A \text{mV}^{-1}$ and the additive voltage bias $V^{layer} = 0.013 \text{V}$.

**Model evaluation.** After the model selection procedure, we use the resulting hyperparameters, using them in new instances of both the software and the physical models. The number of epochs is 100 and the number of evaluations is 100.

**Data availability**
The data supporting the figures of this study are publicly available in the Zenodo repository: https://doi.org/10.5281/zenodo.7956045.