Electronic Supplementary Material (ESI) for Nanoscale Horizons.
This journal is © The Royal Society of Chemistry 2023

Supplementary Information

Adapted MLP-Mixer network based on crossbar arrays of fast and multilevel switching (Co-Fe-B)$_x$(LiNbO$_3$)$_{100-x}$ nanocomposite memristors

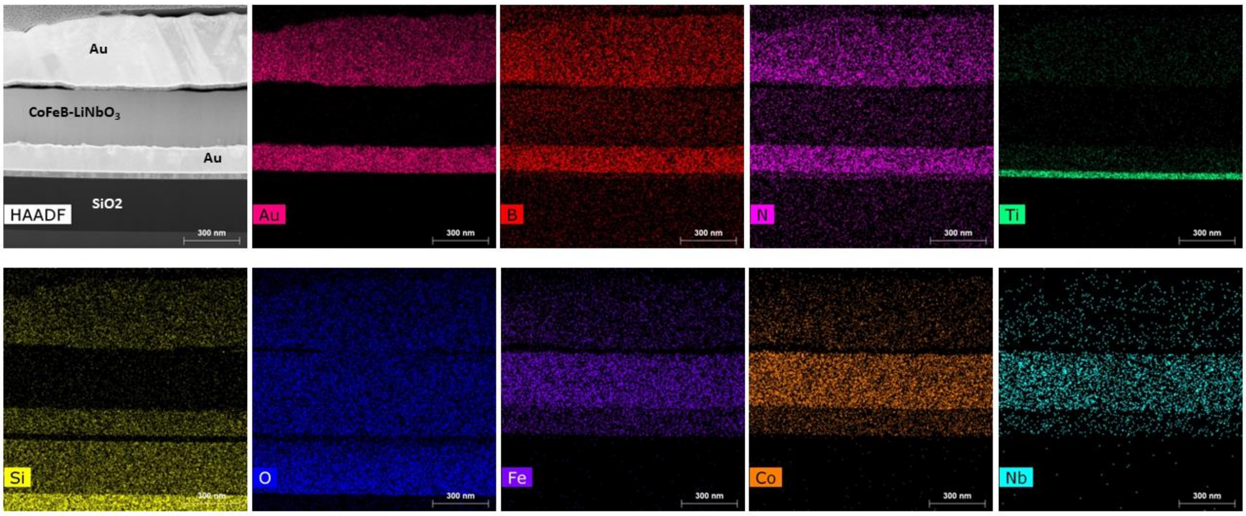


Figure S1. EDX analysis of a single memristor from crossbar array

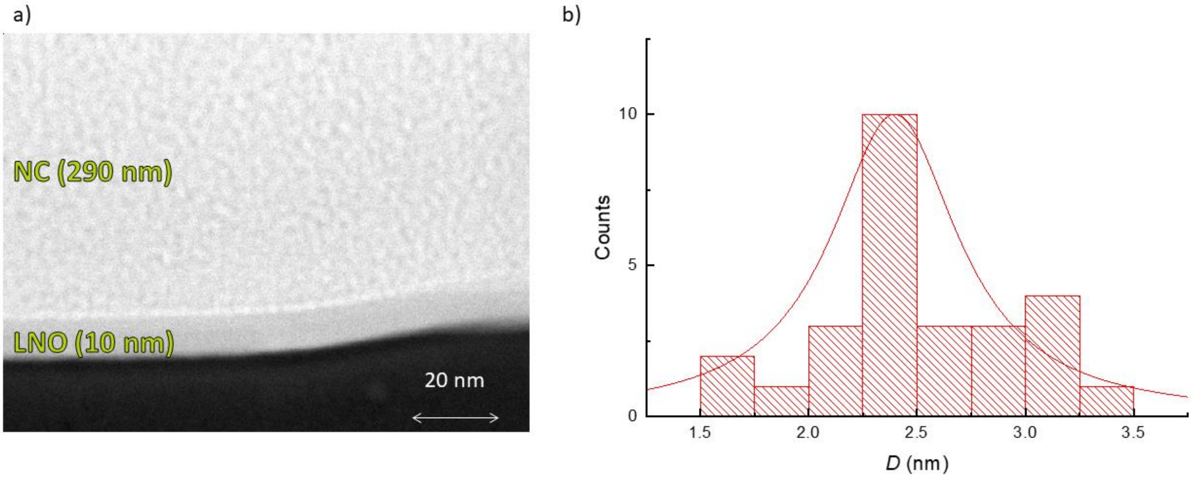


Figure S2. a) Bright field TEM image of active layer of LNO NC memristor; b) distribution of sizes of metal particles in NC layer.

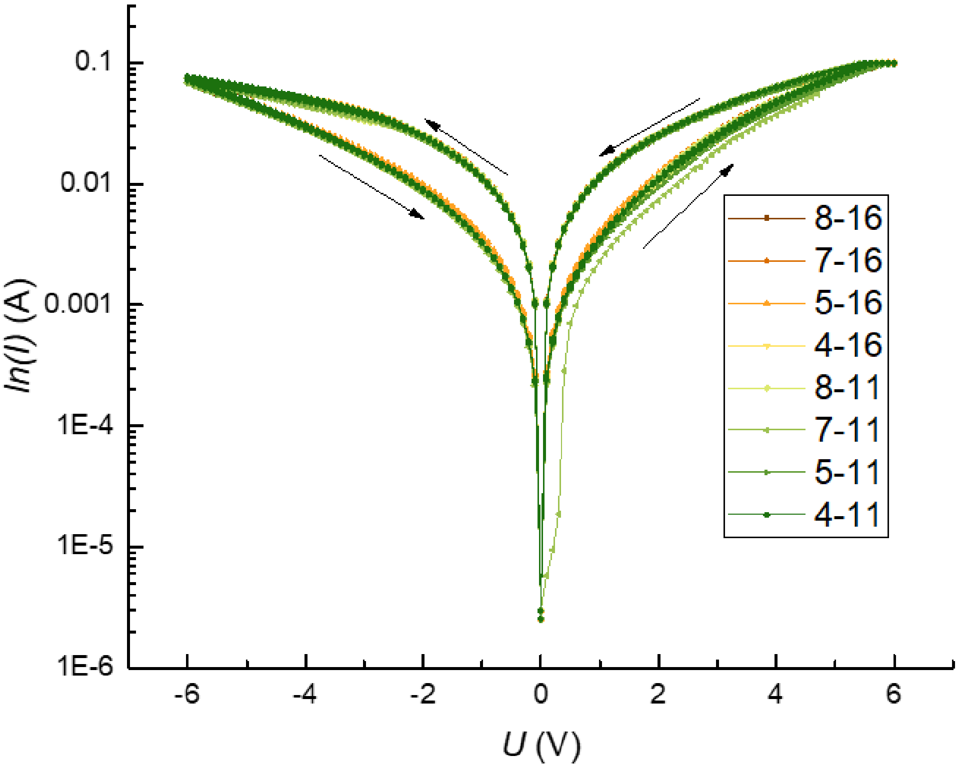


Figure S3. Voltage-current characteristic of eight memristors from the 16x16 crossbar array in logarithmic scale.

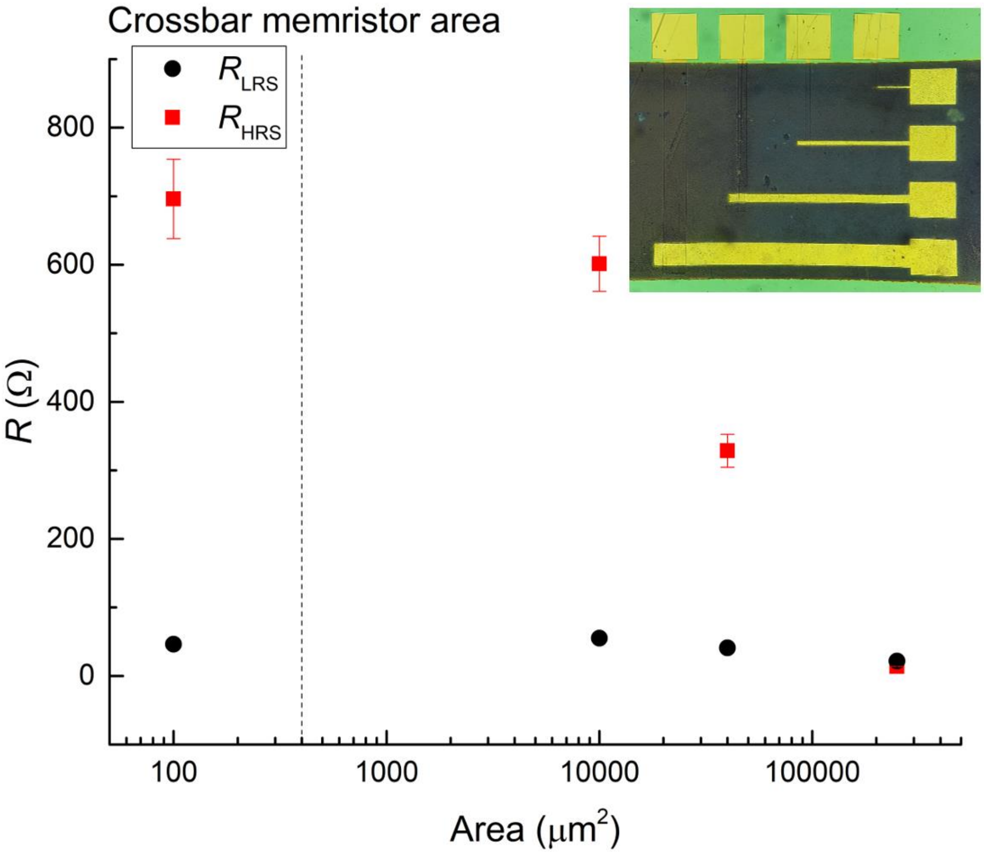


Figure S4. Resistances R_{HRS} and R_{LRS} of cross-point memristors with different square areas. Inset shows microphotograph of cross-point structures with different area.

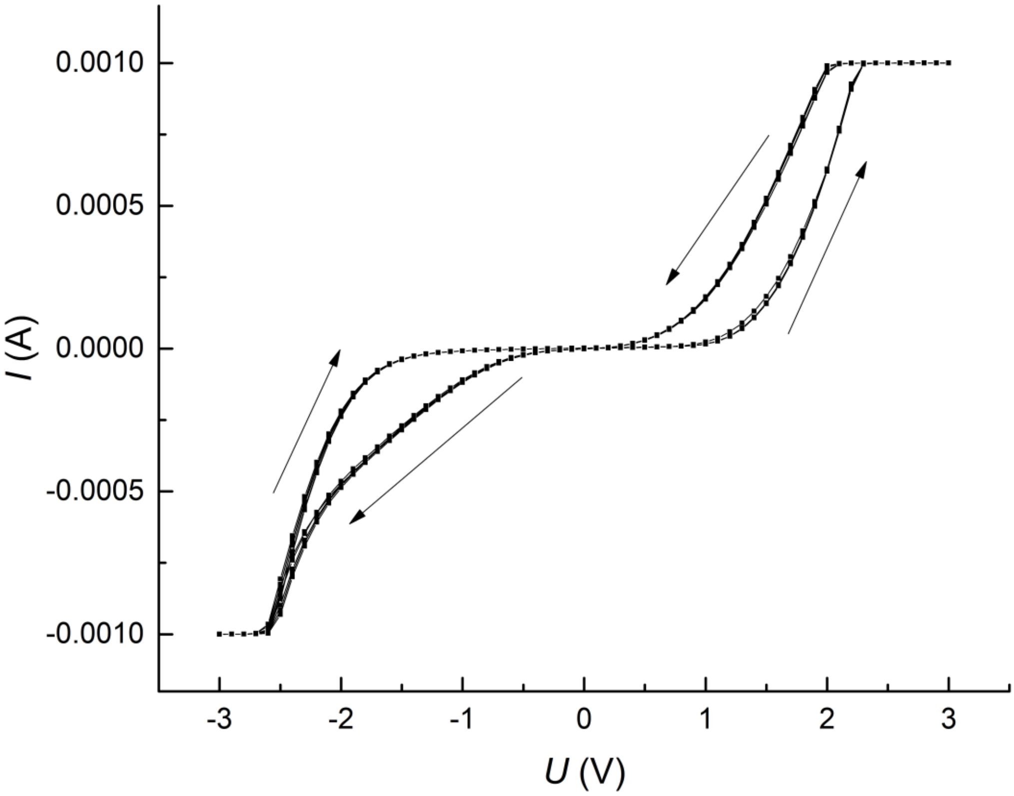


Figure S5. IV-characteristics of single memristor made of the same as in this work active materials with different thicknesses: ~230 nm of NC and ~20 nm $LiNbO_3$

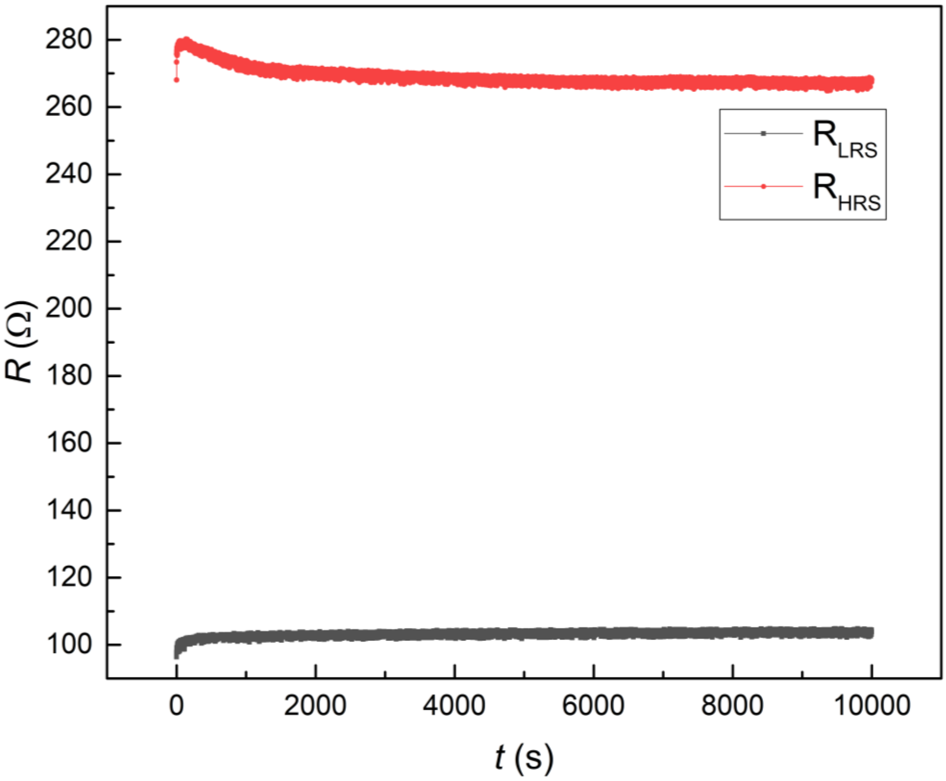


Figure S6. Long retention of CFB-LNO NC memristor.

Supplementary Note 1.

The problem with sneak current estimation is important for crossbar arrays of memristive structures.[1] So, the sneak-path current for the crossbar memristive structures should be estimated. Different cases are considered: crossbars $2 \times 2, 3 \times 3, N \times N$ with and without bus resistance.

For the $N \times N$ crossbar without bus resistance,[1] the total resistance of the crossbar is:

$$R_{tot} = \frac{\tilde{R}\, R_{mem}}{R_{mem} + \tilde{R}},$$

where

$$\tilde{R} = \frac{2\,R}{N-1} + \frac{R}{(N-1)^2}.$$

R_{mem} — the resistance of the memristor being measured, R — resistance of all the other memristors, the same for every memristor. For sneak current

$$I_{sneak} = \frac{U_0}{\tilde{R}}.$$

It is easy to check these formulas explicitly for the cases 2×2, 3×3, if needed. It is possible to obtain this result for the 3×3 case with the equivalent scheme.

In contrast, it is difficult to perform these calculations analytically in general case for the model with bus resistance and determine the exact answer. In this case, the full system of the Kirchhoff's circuit laws should be solved with Gaussian elimination. Special programs such as Cadence Virtuoso can be used for this purpose.[2] Still, it is possible to estimate the right answer.

The following notation will be used: q — full horizontal bus (top bus of our crossbar) resistance, r — full vertical bus (bottom bus of our crossbar) resistance, the current through the chosen memristor will be denoted as $I_{mem,2}$.

The system with the resistance can be transformed in the following way. Let us consider the new system with memristors with $I_{mem,3}$ and resistance $R_{off} + q + r$. In this system

$$I_{mem,1} = U_0/R_{off},\ I_{mem,3} = \frac{U_0}{R_{off}+r+q}\ ,$$

where r and q are bus resistances.

$$I_{mem,2} = \frac{U_0 - i_1 \cdot \frac{r}{N-1} - i_2 \cdot \frac{r}{N-1} - \ldots - i_{N-1}\frac{r}{N-1} - j_1 \cdot \frac{q}{N-1} - j_2 \cdot \frac{q}{N-1} - \ldots - j_{N-1}\frac{q}{N-1}}{R_{off}} \geq$$

$$\geq \frac{U_0 - i_{max}r - j_{max}q}{R_{off}}$$

The hypothesis is that the right ratio of the memristor current to the total (sneak) current I_{mem}/I_{tot} can be estimated as

$$\frac{I_{mem,3}}{I_{tot,3}} \leq \frac{I_{mem,2}}{I_{tot,2}} \leq \frac{I_{mem,1}}{I_{tot,1}}$$

The calculations for the LNO NC crossbar under study are given further.

The worst case

Thus, if the resistance of the buses is neglected, the answer will be as follows for $R_{on} = 100\ \Omega$, $R_{off} = 300\ \Omega$ (one memristor in the upper right corner of the crossbar is in $R_{off} = 300\ \Omega$ and the other memristors are in $R_{on} = 100\ \Omega$, voltage is applied to the uppermost and rightmost buses). The ratio of the current passed through the memristor to the total current:

$$\frac{I_{mem}}{I_{tot}} = \frac{R_{tot}}{R_{off}}\ ,$$

$$\frac{I_{mem}}{I_{tot}} = 0.5 \qquad (N = 2),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.3 \qquad (N = 3),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.05 \qquad (N = 16).$$

The estimation in the case, where the resistance of the buses is not neglected, $R_{on} = 100\ \Omega$, $R_{off} = 300\ \Omega$, $q = 20\ \Omega$, $r = 58\ \Omega$ (Fig. S5):

$$\frac{I_{mem}}{I_{tot}} = \frac{R_{tot}}{R_{off}+q+r}\ ,$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.4 \qquad (N = 2), \qquad \text{exact value } \frac{I_{mem}}{I_{tot}} = 0.5 \qquad (N = 2),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.25 \qquad (N = 3),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.035 \qquad (N = 16).$$

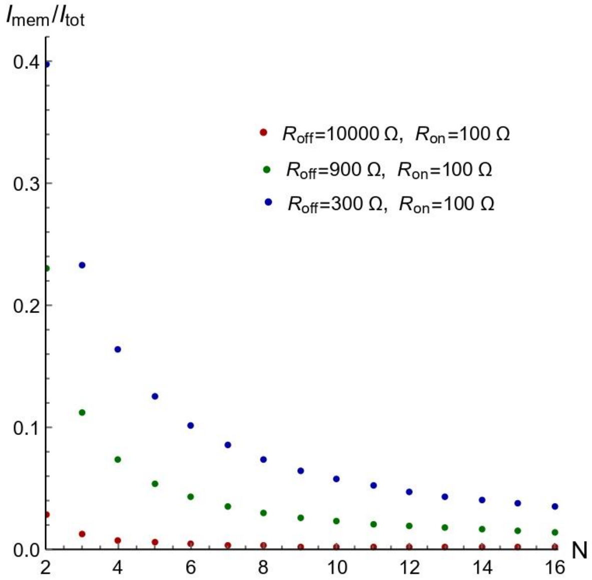


Figure S7. The worst case parametric analysis of the I_{mem}/I_{tot} dependence on the crossbar size N and different R_{off}; bus resistances q = 20 Ω, r = 58 Ω are considered.

The best case

In this case for the $N \times N$ crossbar without bus resistance, the total resistance of the crossbar is

$$R_{tot} = \frac{\tilde{R}\, R_{on}}{R_{on}+\tilde{R}} \,,$$

where

$$\tilde{R} = \frac{2\, R_{off}}{N-1} + \frac{R_{off}}{(N-1)^2} \,.$$

$$I_{sneak} = \frac{U_0}{\tilde{R}} \,.$$

If the resistance of the buses is neglected, $R_{on} = 100\ \Omega$, $R_{off} = 300\ \Omega$, the ratio of the current passing through the memristor to the total current will be:

$$\frac{I_{mem}}{I_{tot}} = \frac{R_{tot}}{R_{on}} \,,$$

$$\frac{I_{mem}}{I_{tot}} = 0.9 \qquad (N = 2),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.79 \qquad (N = 3),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.29 \qquad (N = 16).$$

The estimations can be produced in the case with the resistance of the buses $R_{on} = 100\ \Omega$, $R_{off} = 300\ \Omega$, $q = 20\ \Omega, r = 58\ \Omega$. In this case for the $N \times N$ crossbar with bus resistance the total resistance of the crossbar is

$$R_{tot} = \frac{\tilde{R}\, R_{on}}{R_{on}+\tilde{R}} \,,$$

where

$$\tilde{R} = \frac{2\,R_{off}}{N-1} + \frac{R_{off}}{(N-1)^2} + q + r\,.$$

For I_{mem}/I_{tot} we can obtain (Fig. S6):

$$\frac{I_{mem}}{I_{tot}} = \frac{R_{tot}}{R_{on}}\,,$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.9072 \qquad (N=2), \quad \text{exact value } \frac{I_{mem}}{I_{tot}} = 0.8426 \qquad (N=2),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.82 \qquad (N=3),$$

$$\frac{I_{mem}}{I_{tot}} \approx 0.544 \qquad (N=16)\,.$$

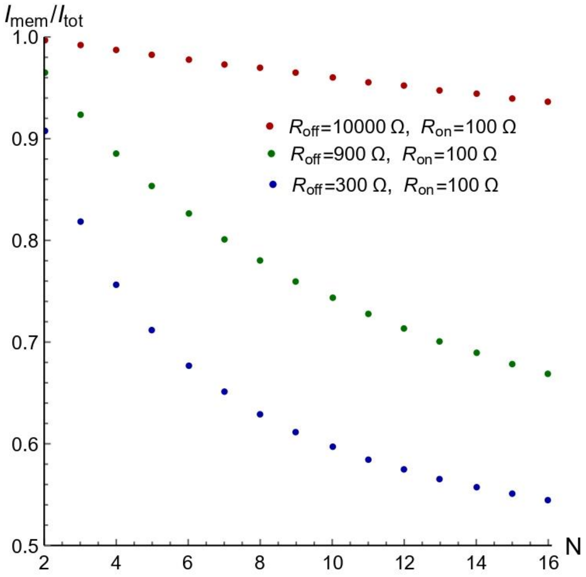


Figure S8. The best case parametric analysis of the I_{mem}/I_{tot} dependence on the crossbar size N and different R_{off}; bus resistances q = 20 Ω, r = 58 Ω are considered.

[1] C. -L. Lo, T. -H. Hou, M. -C. Chen and J. -J. Huang, "Dependence of Read Margin on Pull-Up Schemes in High-Density One Selector–One Resistor Crossbar Array," in IEEE Transactions on Electron Devices, vol. 60, no. 1, pp. 420-426, Jan. 2013, doi: 10.1109/TED.2012.2225147.

[2] Tang, Z.; Wang, Y.; Chi, Y.; Fang, L. Comprehensive Sensing Current Analysis and Its Guideline for the Worst-Case Scenario of RRAM Read Operation. Electronics 2018, 7, 224. https://doi.org/10.3390/electronics7100224

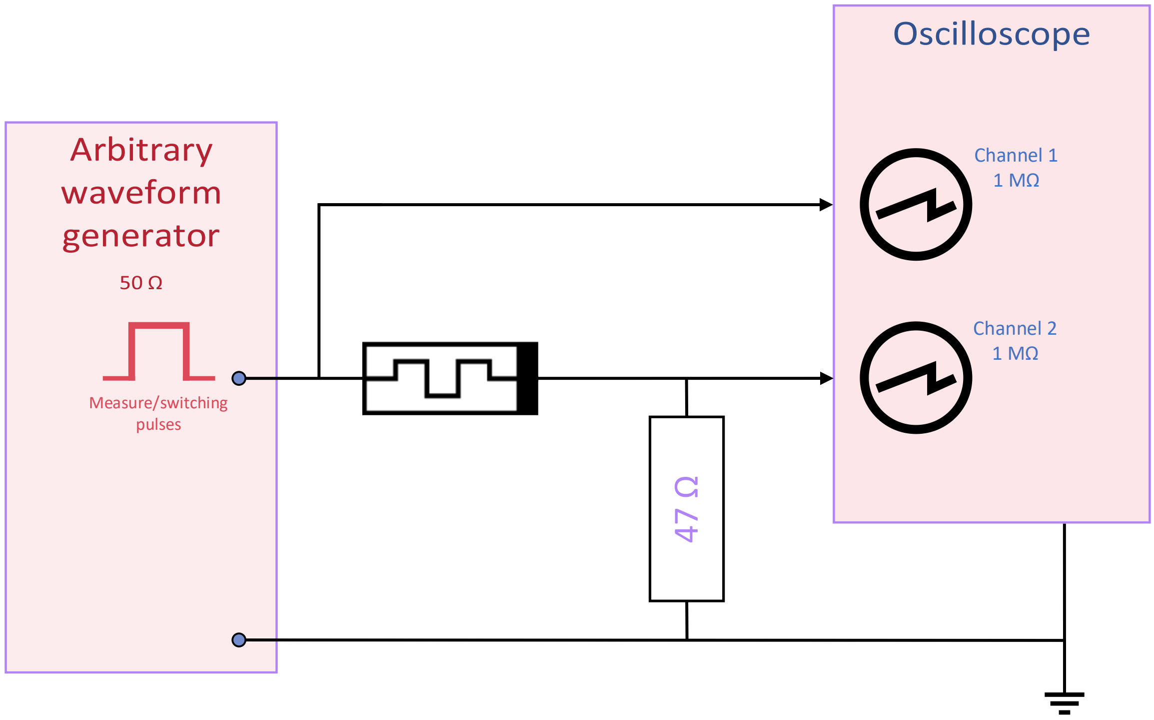


Figure S9. Scheme of the electric circuit, used in switching kinetics experiment.

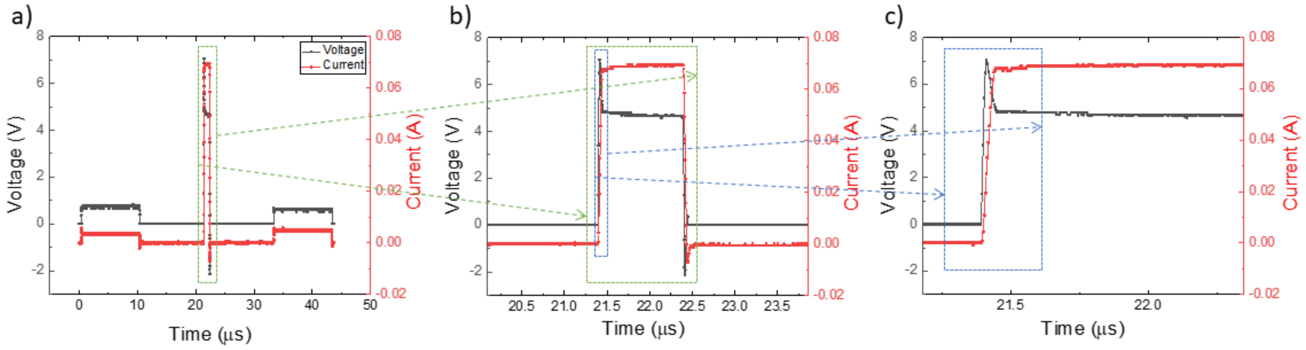


Figure S10. Voltage and current signals during studying of memristor switching kinetic. a) two read pulses and switching pulse between them; b) enlarged view of switching pulse in a); c) enlarged view on the first part of switching pulse.

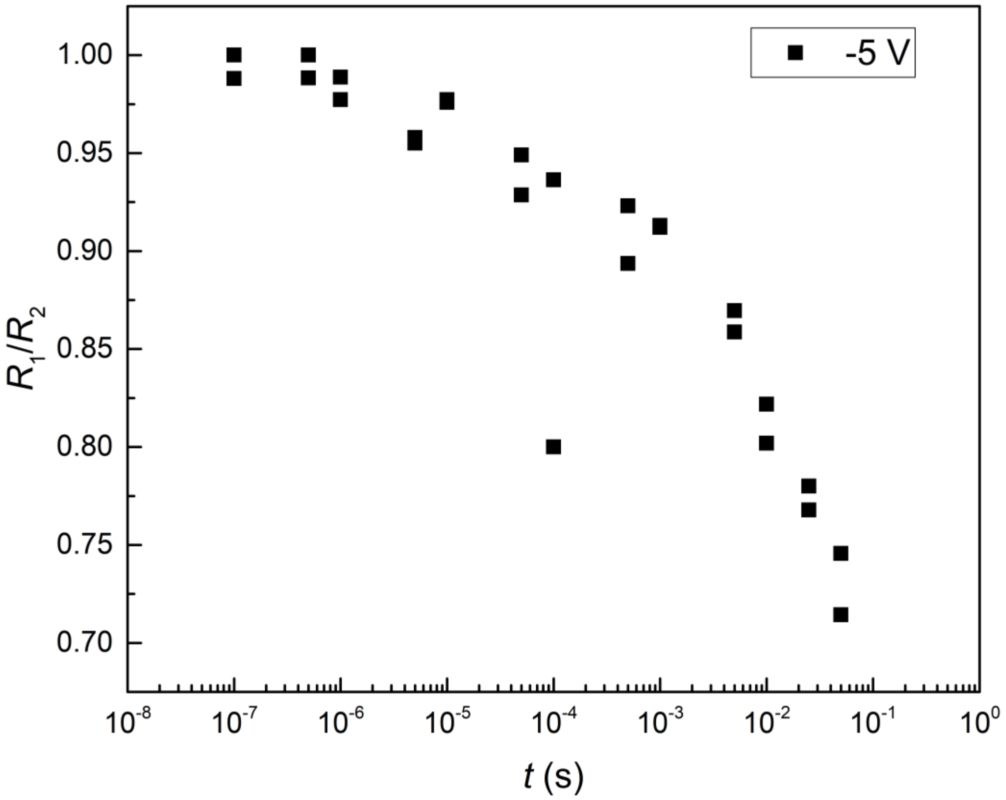


Figure S11. Dependence of initial (R1) to final (R2) resistances ratio vs duration of switching pulse for RESET process

Supplementary Note 2.

The weights for the hardware implementation were obtained via off-chip training. First, a software 4x2 fully connected network was trained for classification of the "0101" and "1010" vectors. When 100% accuracy was achieved, the software full precision weights of the network were obtained (Fig. S12a). Then they were binarized (Fig. S12b, the binarization threshold equaled 0). The binarization process led to no accuracy decrease, which means that binarized weights are sufficient for this simple classification task. Therefore, the binarized weights were transferred to the memristive crossbar as follows: “1” meant that memristor had to be in low-resistance state, while “-1” – high-resistance state.

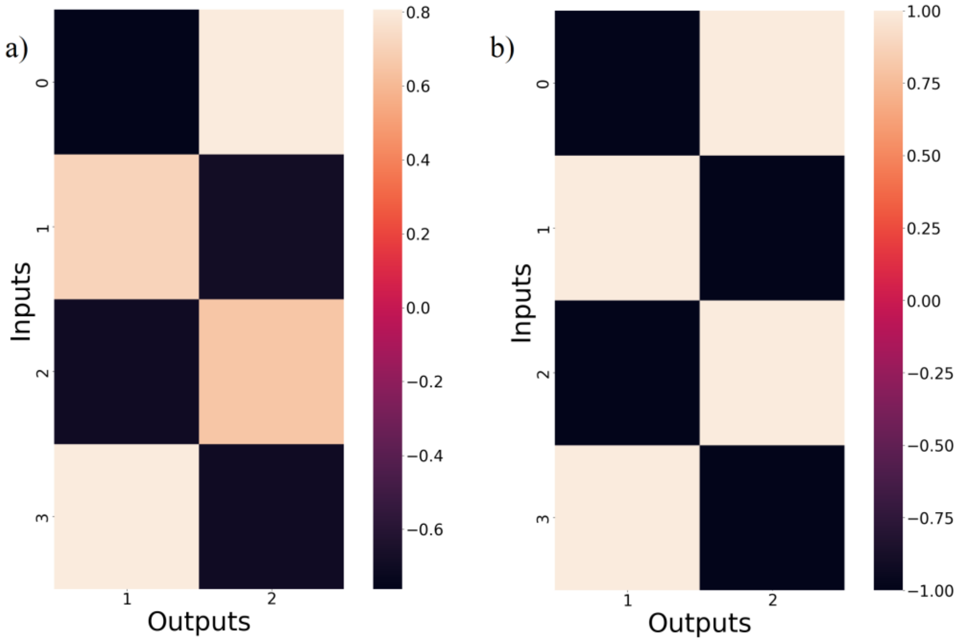
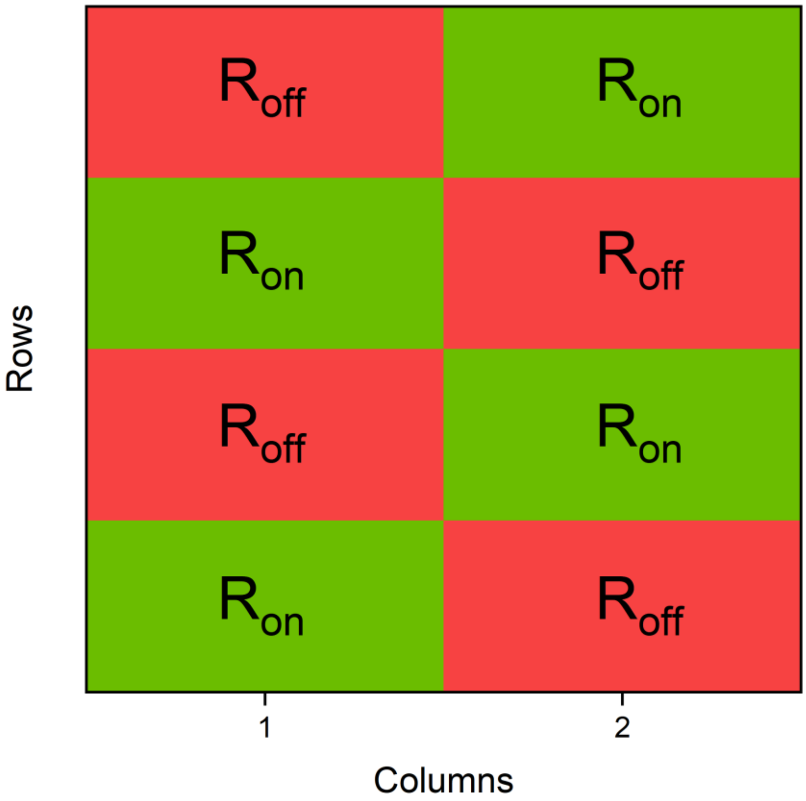


Figure S12. a) The trained software full precision weights of the neural network. b) Binarized weights that were subsequently transferred to the hardware system.



1	2
R_{off}	R_{on}
R_{on}	R_{off}
R_{off}	R_{on}
R_{on}	R_{off}

Figure S13. Weight map of trained hardware perceptron.

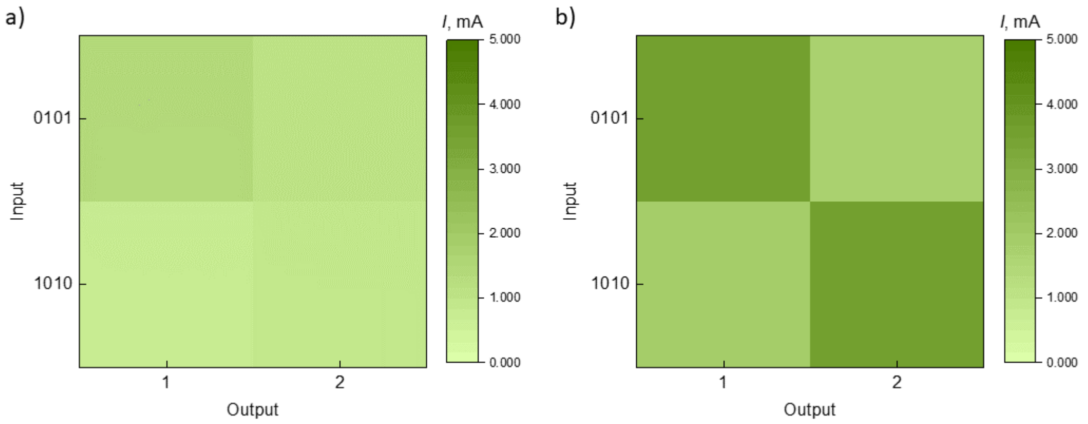


Figure S14. Colormap of output currents for two ideal input images: a) before NCS training b) after training. Note that color bars are the same in a) and b).

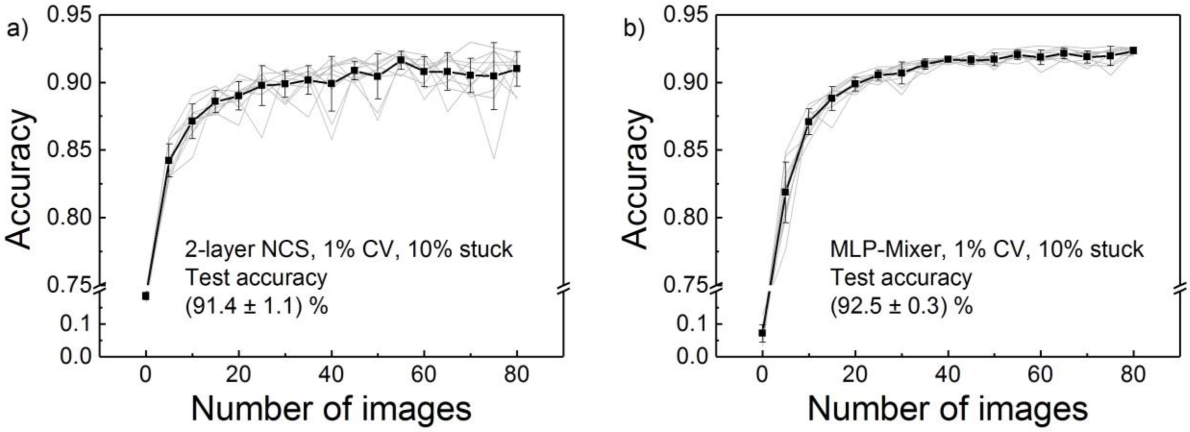


Figure S15. Training curves for the a) 2-layer NCS, b) MLP-Mixer. The experimental CV was utilized for the simulation, 10% of the memristors were stuck in the R_{on} state, i.e., their weights equaled 1.

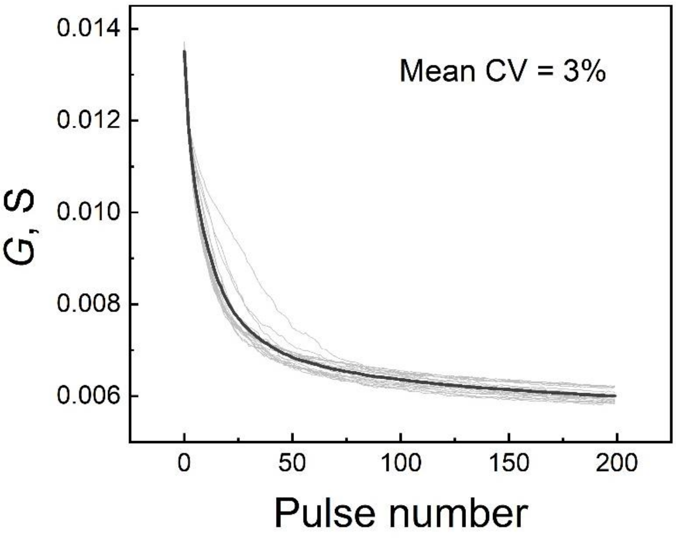


Figure S16. Depression curves obtained from 2 memristive devices (10 curves for each). The averaged curve is marked with a bold line.

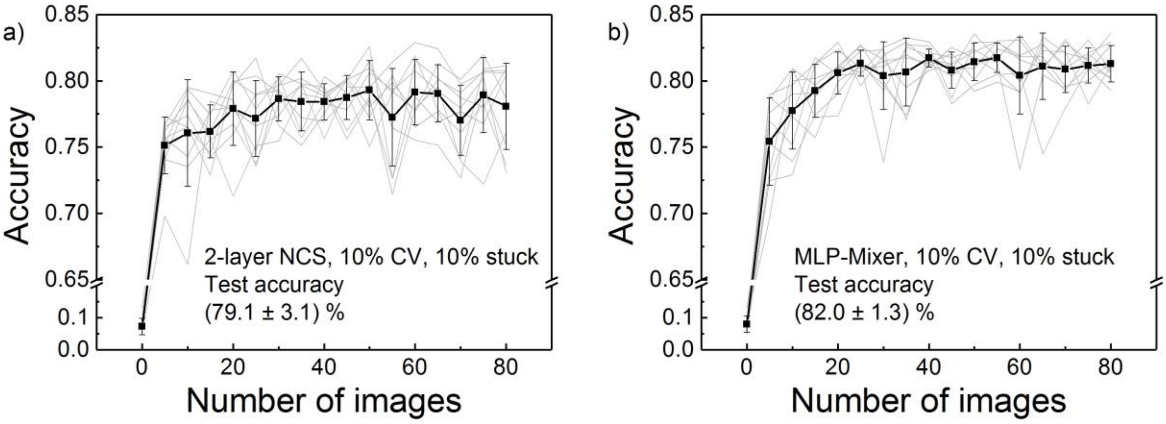


Figure S17. Training curves for the a) 2-layer NCS, b) MLP-Mixer. The 10 % CV was utilized to simulate memristors that are more defective than the NC LNO memristors, 10% of the memristors were stuck in the R_{on} state, i.e., their weights equaled 1.

Table S1. The classification accuracy for the cropped and resized MNIST dataset. The utilized architecture, CV and percentage of the stuck memristors are presented in the table.

Utilized architecture, CV and percentage of the stuck memristors	Mean accuracy ± standard deviation, %	Minimum accuracy, %	Maximum accuracy, %
2-layer NCS, 1% CV, 10% stuck memristors	91.4 ± 1.1	89.3	92.5
MLP-Mixer, 1% CV, 10% stuck memristors	92.5 ± 0.3	91.9	92.9
2-layer NCS, 10% CV, 10% stuck memristors	79.1 ± 3.1	74.1	82.3
MLP-Mixer, 10% CV, 10% stuck memristors	82.0 ± 1.3	80.1	84.1