Electronic Supplementary Material (ESI) for Materials Advances.

Supporting Information

Experimental Method

p-STT MTJ neuron device

The p-STT MTJ neuron device was based on double MgO based perpendicular-magnetic-tunnel-junction (p-MTJ) spin-valve which was deposited using a 12-inch SiO_2 wafer multi-chamber cluster-magnetron sputtering system under a strong vacuum of less than 1×10^{-8} Torr, as shown in Figure 1a. In particular, p-STT MTJ neuron device was vertically stacked on a TiN electrode/Ta buffer/Pt seed/$[Co(0.47\ nm)/Pt(0.23\ nm)]_6$ lower SyAF/spacer/$[Co(0.47\ nm)/Pt(0.23\ nm)]_3$ upper SyAF layer/W bridge layer(0.22 nm)/$Co_2Fe_6B_2$ pinned layer(1.27 nm)/MgO tunneling barrier(1.15 nm)/Fe $Co_2Fe_6B_2$ lower free layer(1.1 nm)/W spacer layer(0.4 nm)/ $Co_2Fe_6B_2$ upper free layer(0.8 nm)/MgO cap(0.8 nm)/capping layer/Ru top electrode(20 nm), and the RF power of the MgO tunneling barrier was varied from 260, 290, 320 and 350 W. Then, the p-MTJ neuron device was ex-situ annealed at 350 ^oC for 30 min under a vacuum below 10^{-6} Torr and a perpendicular magnetic field of 3 Tesla.

Tunneling magnetoresistance (TMR) measurement

The TMR ratio were estimated for all 12-inch wafers by using the current-in-plane tunneling (CIPT) technique at room temperature. Measurement was carried out along the notch direction for every 2 cm.

Grain size distribution and MgO crystallinity

To investigate the grain size distribution of the p-STT MTJ neuron device with different sputtering power of the MgO tunneling barrier (260, 350 W), the samples were cut into 1 × 1 cm^2 and TEM samples were prepared with focused ion beam (FIB). The crystallinity and grain size distribution of the MgO tunneling barrier depending on the RF sputtering power was examined by using high-resolution transmission-electron-microscopy (HR-TEM) at 200 keV, as shown in Figure 2. In addition, to investigate properties of bulk MgO film such as crystallinity and surface roughness, MgO film was stacked on Si substrate / MgO (10 nm, RF power 260, 290, 320 and 350W) and was ex-situ annealed at 350 ^oC for 30 min. and Supplementary 3. The optical properties to compare the defect state density and the energy gap of MgO film depending on sputtering power was observed by using spectroscopic ellipsometry (SE, J. A. Woollam Co. VASE Ellipsometer), as shown in Supplementary 4.

Integrate behavior measurement

The wafers were cut into 2 × 2 cm^2 pieces and it were patterned into p-MTJ with a device size of 1.6 × 1.6 μm^2 using ion milling and photolithography. Then R-V curve, switching current and integration behavior of p-MTJ neuron devices were measured using Keithley 236 source measure unit and an Agilent B2902A semiconductor parameter analyzer

SNN simulation method

In generals, gradient based learning algorithm such as gradient decent method with backpropagation cannot be used to train SNN because the spike signals were non-differentiable function. Thus, the STDP algorithm has been used to train SNN. STDP is a biological-plausible

unsupervised learning method, in which synaptic weights between two adjacent neurons were updated depending on spike timing difference between two neurons. Unlike global learning algorithms such as backpropagation that use both local and up-stream gradient to update synaptic weight, the STDP learning algorithms use only local information between spiking neurons. This makes learning of neural network somewhat simple, but it causes less accuracy. Several learning techniques can be used to increase the accuracy of single-layer SNNs such as winner-takes-all (WTA), and homeostasis, etc. In this work, a simplified STDP learning rule with winner-takes-all (WTA) was used to train single-layer SNN, as shown in Figure S8b. In this STDP learning, the MNIST images are processed by clustering and labeling, so that a specific output neuron of the network fires only a specific input number. Since the STDP learning method used in this simulation is an unsupervised learning method that performs only clustering. During training, the mode table of each output neuron was made, and the accuracy was calculated by comparing this table with the label during the accuracy test.

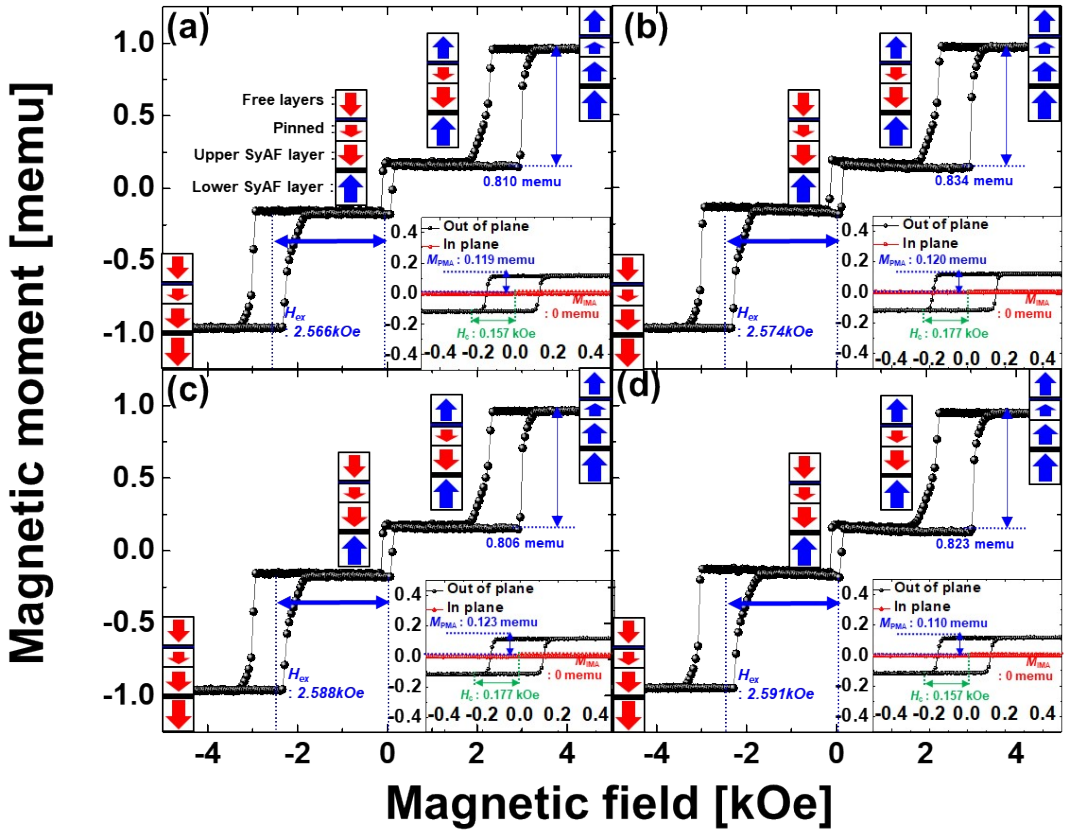

Fig. S1. Static ferro-magnetic properties of the p-STT MTJ neuron device depending on MgO sputtering in the applied perpendicular magnetic field (i.e., -4 ~ + 4 kOe). (a) MgO RF sputtering power 260, (b) 290, (c) 320, and (d) 350 W. The static ferro-magnetic properties of the p-STT MTJ neuron devices depending on sputtering power of the MgO tunneling barrier were measured with vibrating sample magnetometer (VSM). The magnetic property presented a similar behavior where the magnetic moment of the $Co_2Fe_6B_2$ free layer were about 120 μemu and coercivity of about 160 Oe for all sputtering power with a good squareness, indicating that all p-STT MTJ neuron device performed a good interfacial perpendicular magnetic anisotropy (i-PMA), as shown in the insets of Figure S1.

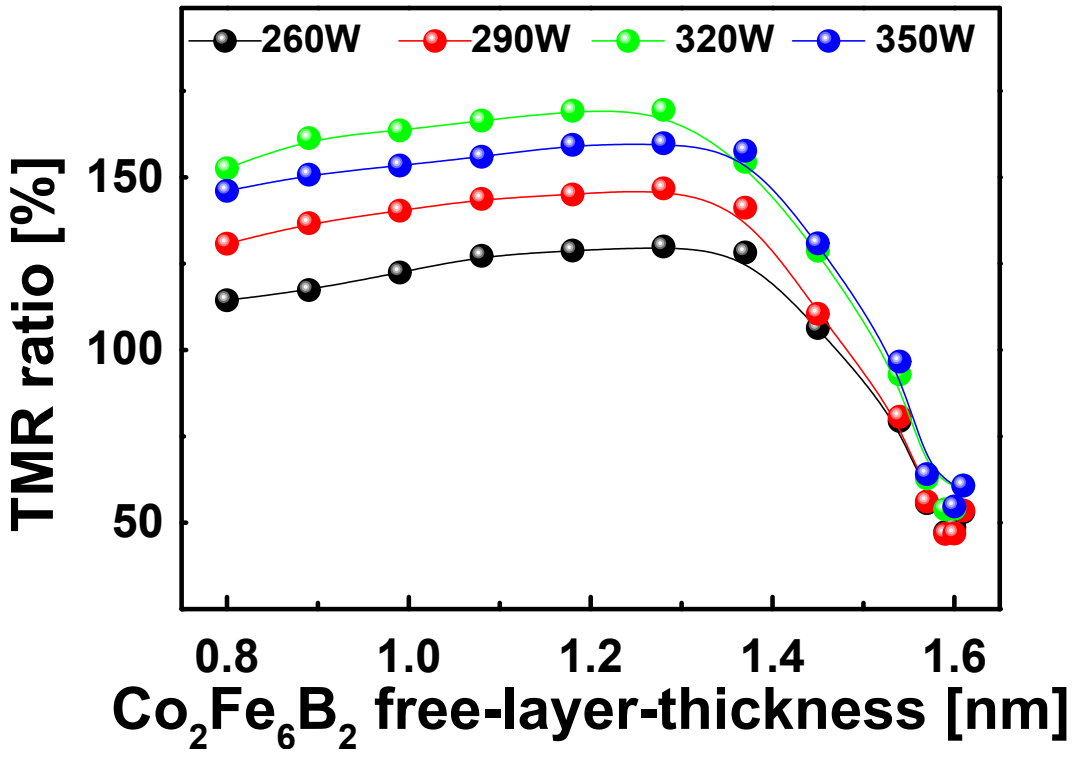

Fig. S2. Dependency of the TMR ratio on the RF sputtering power and Co2Fe6B2 free-layer thickness of Co2Fe6B2/MgO based p-STT MTJ neuron device on 12-inch wafers. The TMR ratio was measured using CIPT of unpatterned p-STT MTJ neuron devices depending on sputtering power of the MgO tunneling barrier. The TMR ratio peaked at certain thickness of $Co_2Fe_6B_2$ free-layer thickness (~1.3 nm). In addition, the TMR increased from 130 % to 170 % as the sputtering power is increased from 260 to 320 W and it decreased to 160 % at sputtering power of 350 W. According to the first principle calculation by Bulter and Mathon [Phys. Rev. B, vol. 63, no. 5 (2001) and Phys. Rev. B, vol. 63, no. 22 (2001)], the high TMR ratio can be achieved by the spin filtering of the Δ_1 coherent tunneling in a crystalline Fe(001)/MgO(001) /Fe(001) MTJ. It was related that the MgO sputtering power of 320 W would have the best f.c.c. crystallinity of the MgO tunneling barrier (001) [1].

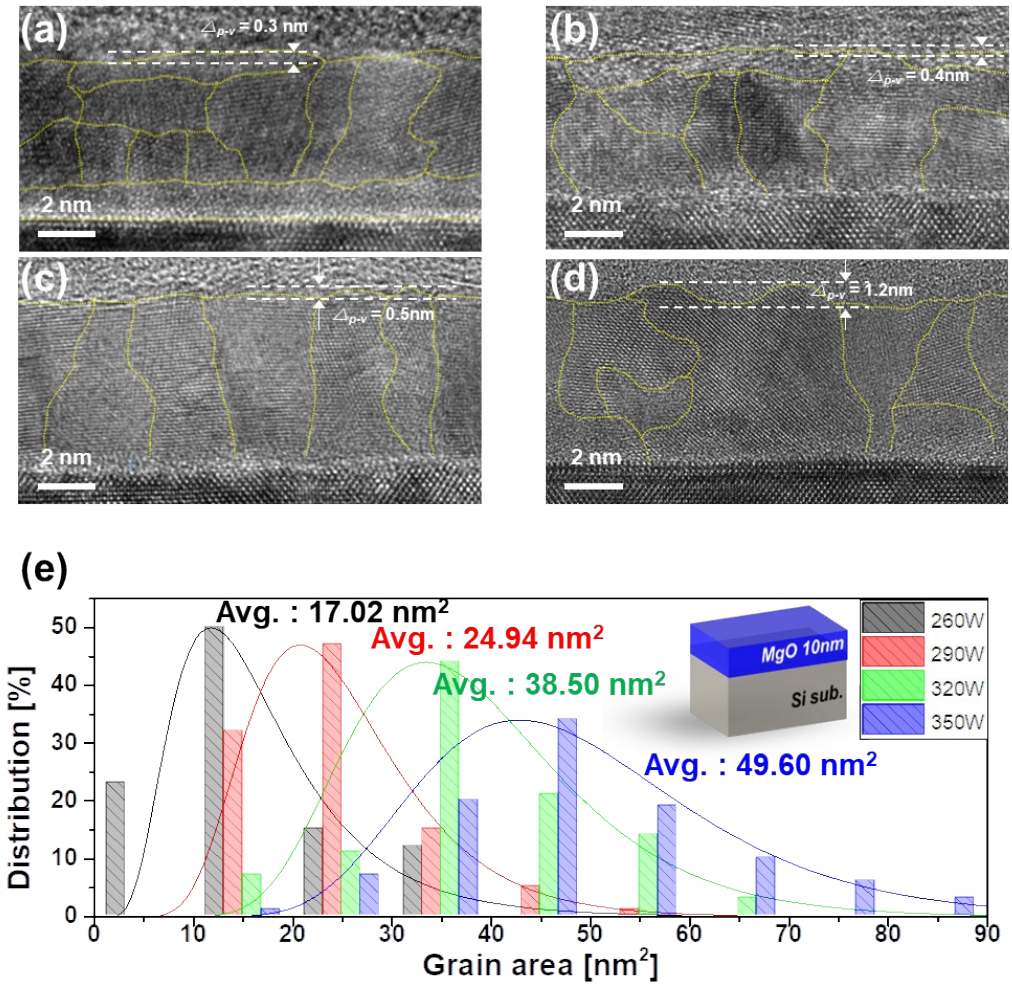

Fig. S3. HR-TEM images of 10-nm thick MgO film depending on the RF sputtering power for ex-situ annealing at 350 oC. (a) 260, (b) 290, (c) 320, (d) 350 W, and (e) grain area distribution. Since it is very difficult to investigate the dependency of the surface roughness on the MgO tunneling barrier sputtering power, 10-nm thick MgO films were sputtered on the Si substrate, and then the surface roughness and the grain area distribution of 10-nm thick MgO films were estimated by observing HR-TEM images. The average grain area of the MgO films increased from 17.02 to 49.6 nm^2 as the sputtering power increased from 260 to 350 W, which was consistent with the grain size distribution of Figure 2. In addition, the surface roughness of the MgO film increased from 0.3 nm to 1.2 nm, as the sputtering power enhanced from 260 to 350 W, reducing the hybridization of Fe_{3d}-O_{2p} and Co_{3d}-O_{2p} at interface between $Co_2Fe_6B_2$ and MgO tunneling barrier [2-4]. Thus, the TMR ratio of p-STT MTJ neuron devices increased when the sputtering power enhanced from 260 to 320 W, since the f.c.c. crystallinity of the MgO tunneling barrier was improved with increasing sputtering power. Otherwise, the sputtering power increased further (i.e. 350 W), the TMR ratio significantly reduced, since the surface roughness of MgO tunneling barrier was enhanced.

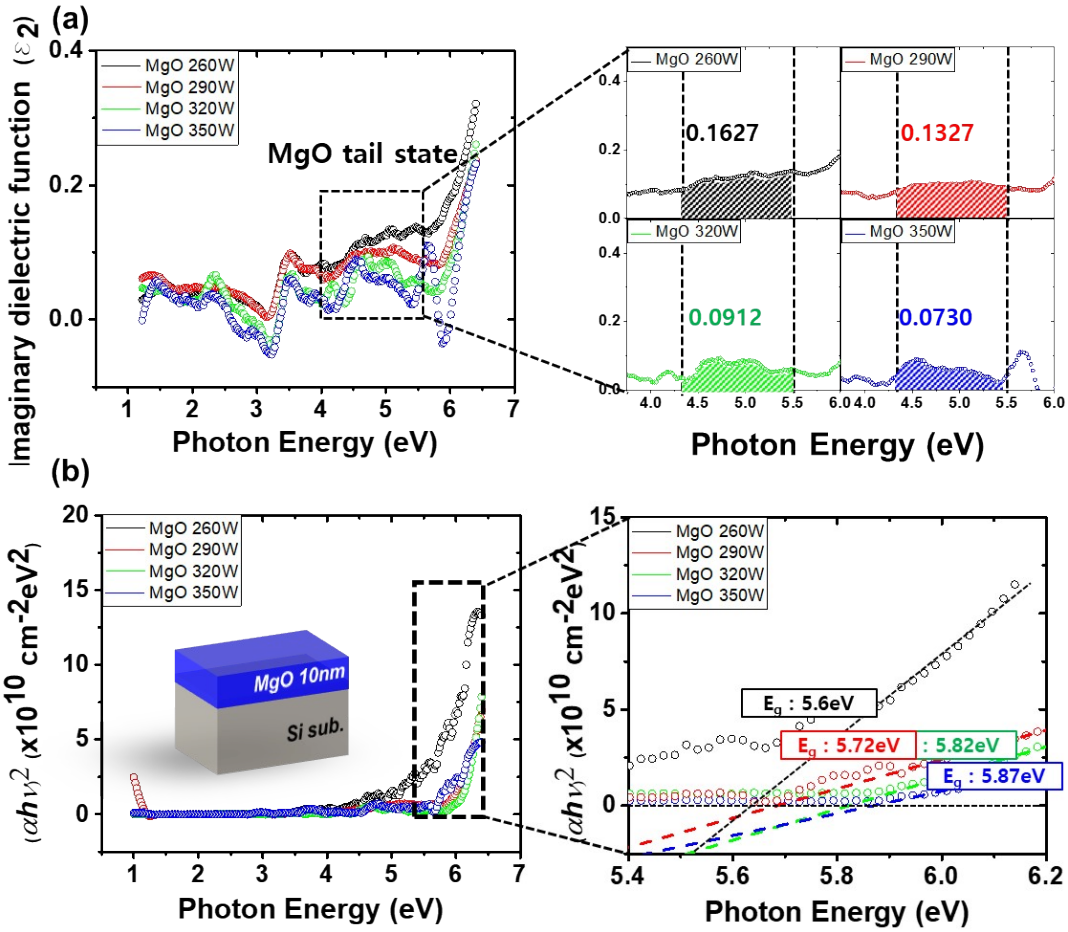

Fig. S4. Dependency of optical properties of 10-nm thick MgO film on the RF sputtering power for ex situ annealing at 350 °C: (a) imaginary dielectric function vs. photon energy between 260, 290, 320 and 350 W and (b) the MgO tunneling barrier energy band gap analyzed by the Tauc method. Since, it is very difficult to observe the detailed f.c.c. crystallinity of the MgO tunneling barrier, the defect states of 10-nm thick MgO thin-films were investigated by measuring the optical properties using spectroscopic ellipsometry (SE), as a function of the MgO sputtering power. The absorption of photons was calculated at the range from 4.3 to 5.5 eV (i.e., MgO tail state) with the imaginary parts (ε_2) of reflected phonons where a higher ε_2 would lead to higher defect state density, as shown in Figure S4a. The defect state density decreases from 0.1627 to 0.0730 as the sputtering power of the MgO tunneling barrier increases from 260 to 350 W. In addition, the absorption coefficients (α) of the MgO tunneling barrier depending on the sputtering power were estimated, as shown in Figure S4b. α is defined by $\alpha(\upsilon) = (h\upsilon - E_g)^n / h\upsilon$, where E_g is the energy bandgap of the MgO tunneling barrier, $h\upsilon$ is the photon energy, and n=0.5 for direct energy bandgap material. By plotting $(\alpha h\upsilon)^2$ vs. photon energy, the energy bandgap can be estimated which increases from 5.6 to 5.87 eV with the sputtering power of the MgO tunneling barrier, as shown in Figure S4b. These results indicate that the f.c.c. crystallinity of the MgO tunneling barrier increases with the MgO sputtering power since a higher sputtering power showed a lower defect state density and a higher energy bandgap of the MgO tunneling barrier.

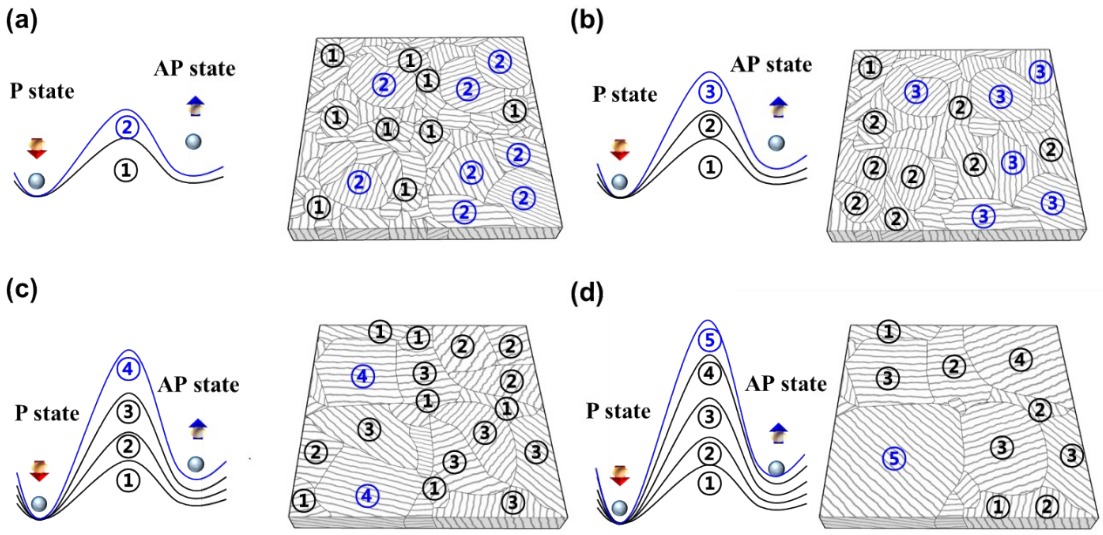

Fig. S5. Schematic of grain size distribution and potential energy distribution vs. location for the MgO tunneling barrier, depending on the average grain size and distribution, where the MgO tunneling barrier thickness of all p-STT MTJ neuron devices were 1.15-nm. (a) 0.418, (b) 0.589, (c) 0.935, and (d) 1.141 nm. In order to understand the integrate mechanism of the p-STT MTJ neuron device depending on the grain size distribution of the MgO tunneling barrier, the schematic concept of energy barrier from parallel to anti-parallel state was reviewed as a function of a grain size distribution. The spin-torque switching energy barrier from parallel to anti-parallel state of the MgO tunneling barrier at the grain boundaries would be much smaller than the grain insides [5, 6]. The integrate behavior of p-STT MTJ neuron devices would be affected by the grain size distribution as a smaller grain size needs less spin-torque energy, as shown in Figure S5a. As consecutive input spikes were applied to the p-STT MTJ neuron devices, a smaller grain size would be spin-torque switched earlier. Thus, the magnitude of the resistance change from parallel to anti-parallel state becomes higher as the grain size of the MgO tunneling barrier was increased from 0.418 to 1.141 nm.

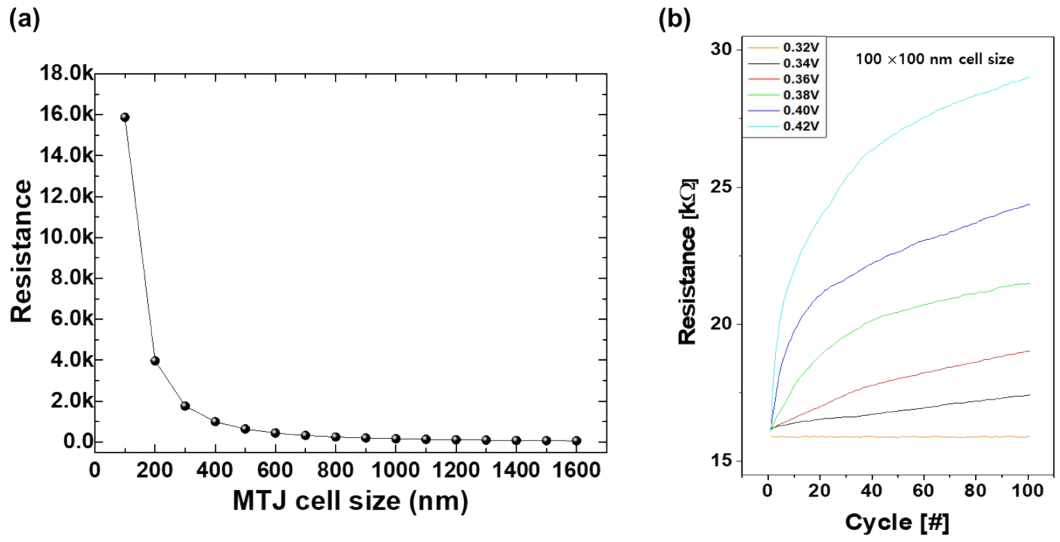

Fig. S6. Projected resistance of p-STT MTJ neuron device scaling-down on p-STT MTJ neuron device size. (a) Dependency of parallel resistance on p-STT MTJ neuron device size, and (b) Simulated integrate behavior of p-STT MTJ neuron device at 100 × 100-nm cell-size depending on spike amplitudes. (i.e., 0.32 ~ 0.42V). The dependency of integrate behavior for the p-STT MTJ neuron device with 100 × 100-nm cell-size showed a similar to that with 1.6 × 1.6 μm cell-size in Figure 3c in main manuscript.

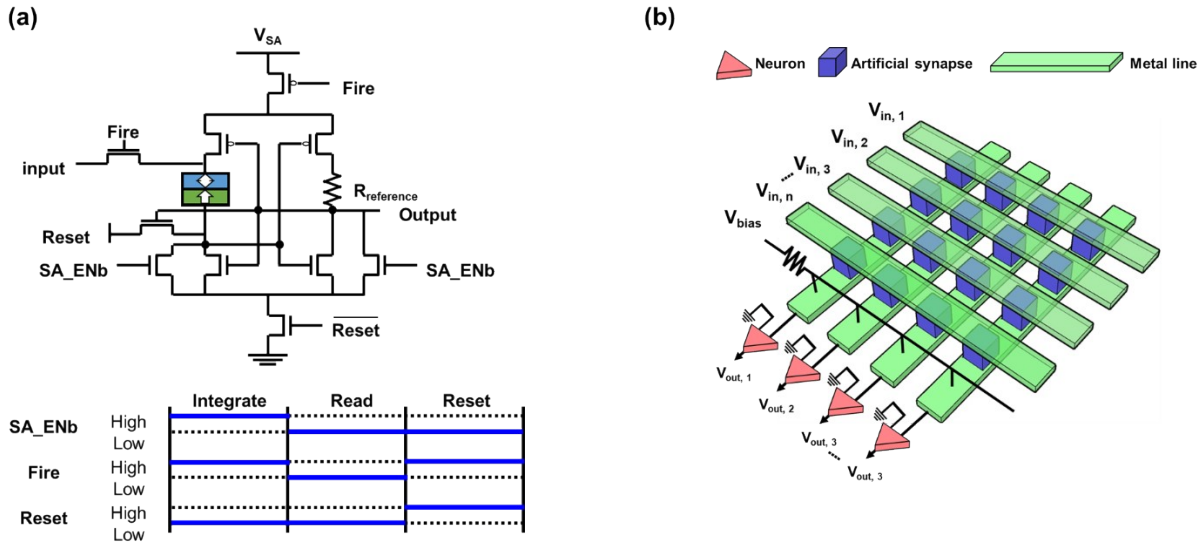

Fig. S7. Spiking neural network. (a) p-STT neuron consisting of 1 p-STT MTJ neuron device, 7 n-MOSFETs, 3 p-MOSFETs and 1 reference resistance (R_{ref}) and SA_ENb, Fire, and Reset pulse to operate a neuron according to input time. (b) schematic of spiking neural network implemented with a crossbar array. The neuron fires when the resistance of the p-STT MTJ neuron device exceeds the reference resistance (R_{ref}), as shown in Figure S7a. The neuron receives signals from the controller to perform integrate, read, and reset operation, as shown in bottom of Figure S7a. For analog matrix-vector multiplication, the cross-point synapse array was connected to the p-STT MTJ neurons, as shown in Figure S7b.

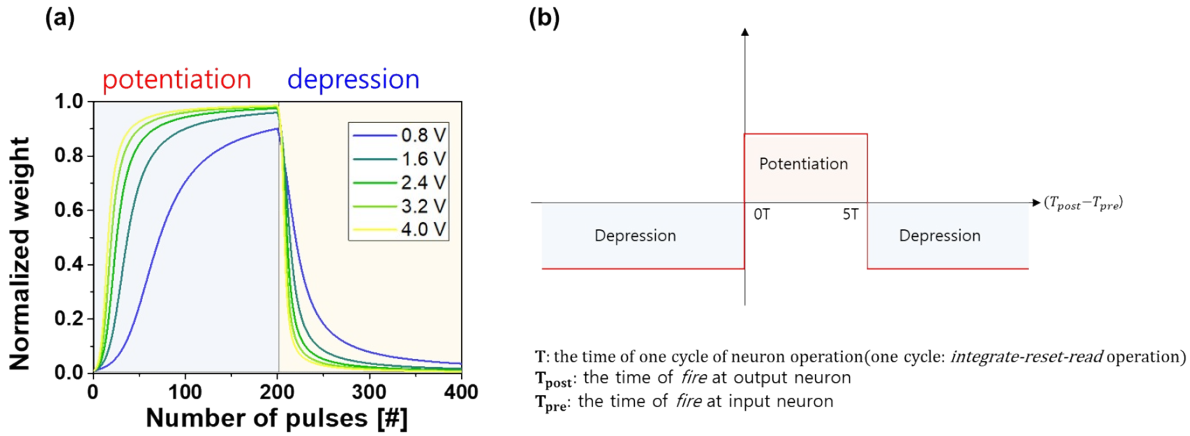

Fig. S8. Synapse characteristics for SNN. (a) Simulated potentiation and depression characteristic of memristor synapse depending on spike amplitude (i.e., 0.8 ~ 4.0 V) (b) simplified STDP learning rule used in simulation. A general potentiation and depression characteristic of memristor was selected [7, 8]. STDP learning rule was used with a simplified potentiation and depression characteristic.

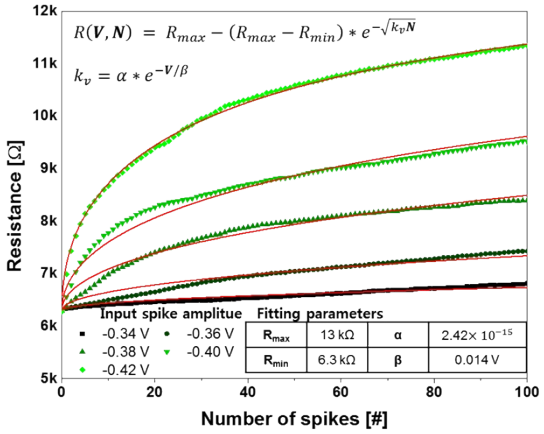

Fig. S9. Fitting of integrate behavior with the resistance vs. number of input spikes for the scale-downed (i.e., 100 × 100-nm cell size) p-STT MTJ neuron devices, depending on input spike amplitude and using Sweibul function. The fitting curve was necessary for the p-STT MTJ neuron simulation.

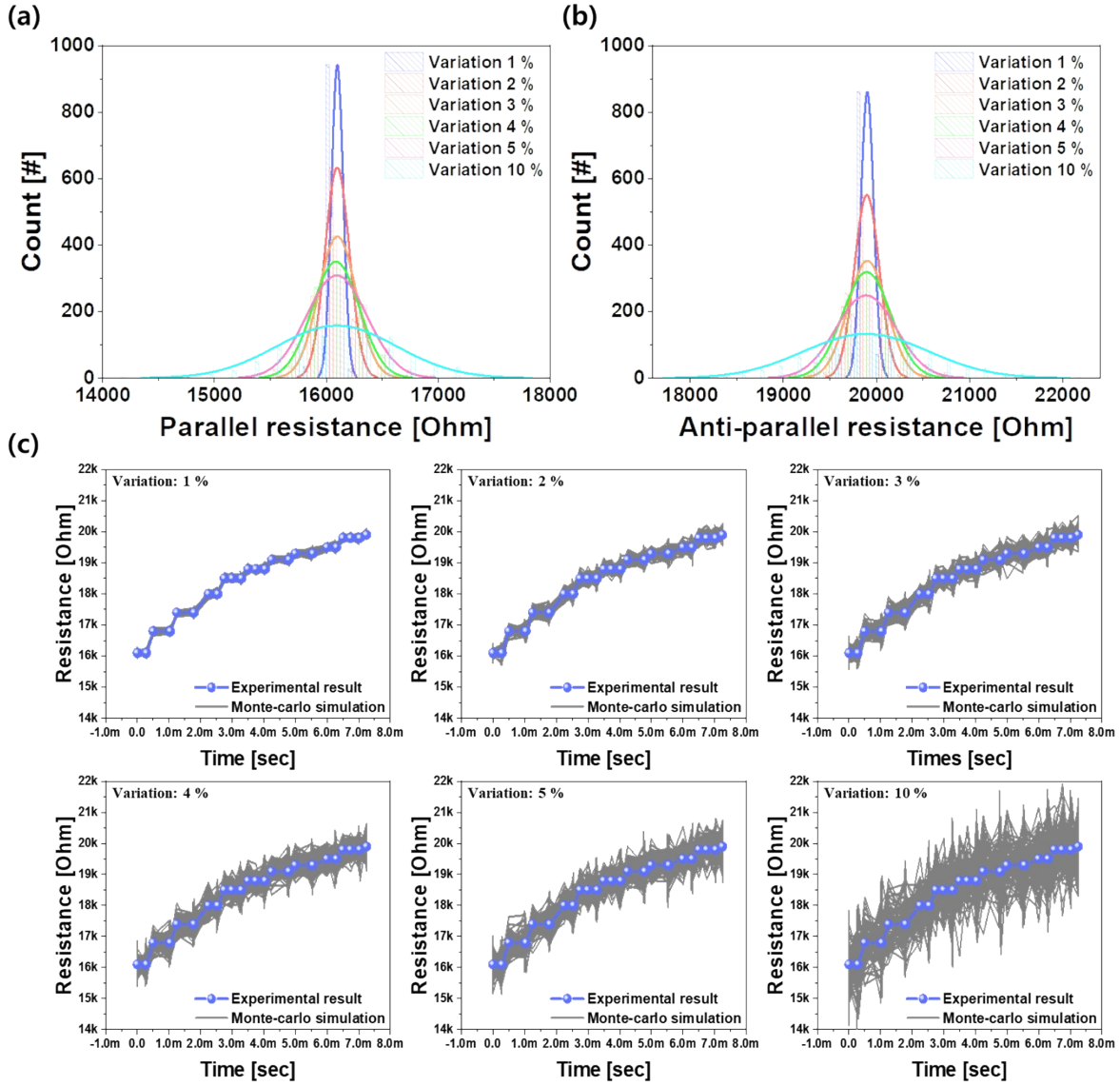

Fig. S10. Monte Carlo simulation of p-STT MTJ neuron device with resistance variation. Resistance distribution of (a) parallel state, (b) anti-parallel state, (c) integrate behaviour depending on the resistance variation.

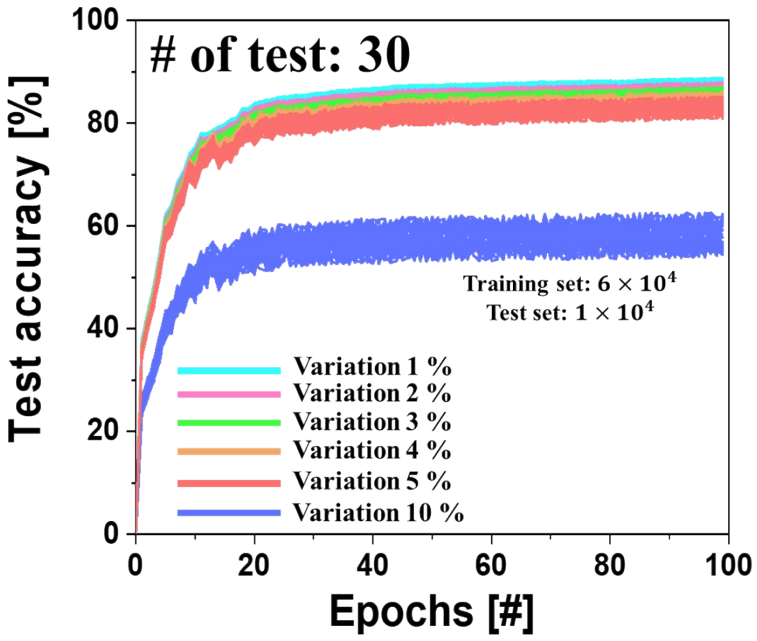

Fig. S11. Pattern recognition accuracy depending on the resistance variation.

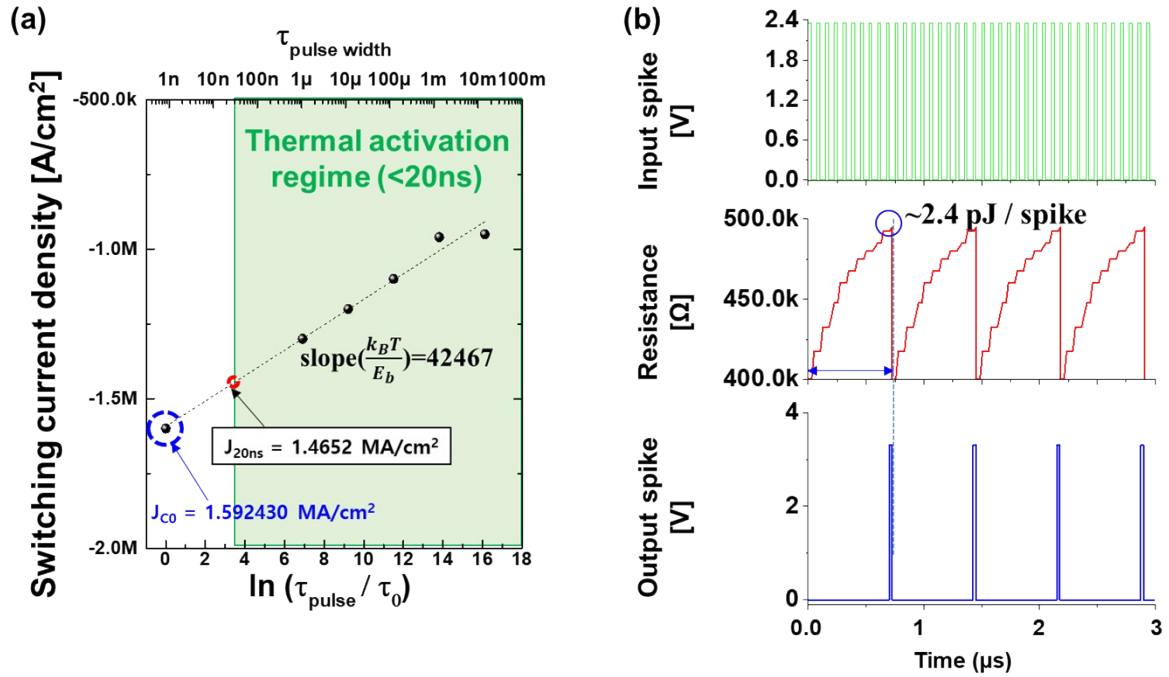

Fig. S12. Power consumption (i.e., pJ/spike) of p-STT MTJ neuron device. (a) switching current density depending on pulse width, where the full spin-torque switching current density with 20 ns switching speed was postulated to be 1.4652 MA/cm2 and (b) input spike, simulated integrated resistance, and output spike as a function of the input spike number of p-STT MTJ neuron device at 20-nm cell size. Maximum power consumption means a spin-torque full switching of the MgO tunneling barrier for a p-STT MTJ neuron device so that the power consumption per spike would be much less than the maximum power consumption. For p-STT MTJ neuron device with average grain size of 0.935 nm, the spin-torque switching current density with a full spin-torque switching of 20 ns was estimated using the thermal activation model of Eq. (1),

$$Thermal\ activation\ model: J_C = J_{C0}\left\{1 - \left(\frac{k_B T}{E_b}\right)\ln\left(\frac{t_{pw}}{\tau_0}\right)\right\} \cdots Eq.\ (1)$$

where, J_{C0} is the switching current density at 0 K, E_b is the switching energy barrier between parallel and anti-parallel states, t_{pw} is the pulse width and τ_0 is the inverse of the attempt frequency (~1 ns). The slope of J_C vs. $\ln(t_{pw}/\tau_0)$ is $k_B T/E_b$, as seen in Figure S12a. Assuming that the integrate behaviour would be the same for cell size of 20-nm of the p-STT MTJ neuron device, the integrate behaviour was estimated as shown in Fig. S12b. The power of the sense amplifier during read operation was calculated using a discrete SPICE model (NMOS: 2N7000, PMOS: BS250) which showed a total power of about 853.6 pW.

Reference

1. Lee, D.-Y.; Seo, H.-T.; Park, J.-G. Effects of the radio-frequency sputtering power of an MgO tunneling barrier on the tunneling magneto-resistance ratio for Co 2 Fe 6 B 2/MgO-based perpendicular-magnetic tunnel junctions. *Journal of Materials Chemistry C* **2016,** 4, (1), 135-141.
2. Chae, K.-S.; Lee, D.-Y.; Shim, T.-H.; Hong, J.-P.; Park, J.-G. Correlation of the structural properties of a Pt seed layer with the perpendicular magnetic anisotropy features of full Heusler-based Co2FeAl/MgO/Co2Fe6B2 junctions via a 12-inch scale Si wafer process. *Applied Physics Letters* **2013,** 103, (16), 162409.
3. Takemura, Y.; Lee, D.-Y.; Lee, S.-E.; Park, J.-G. Dependency of tunneling magnetoresistance ratio on Pt seed-layer thickness for double MgO perpendicular magnetic tunneling junction spin-valves with a top Co2Fe6B2 free layer ex-situ annealed at 400 C. *Nanotechnology* **2016,** 27, (48), 485203.
4. Choi, J.-Y.; Lee, D.-g.; Baek, J.-U.; Park, J.-G. Double MgO-based perpendicular magnetic-tunnel-junction spin-valve structure with a top co 2 Fe 6 B 2 free layer using a single SyAF [co/Pt] n layer. *Scientific reports* **2018,** 8, (1), 1-7.
5. MacLaren, J.; Willoughby, S. First principles determination of the effects of a grain boundary on the anisotropy and exchange energies in Co 1 Pt 5 superlattices. *Journal of Applied Physics* **2001,** 89, (11), 6895-6897.
6. Victora, R.; Willoughby, S.; MacLaren, J.; Xue, J. Effects of grain boundaries on magnetic properties of recording media. *IEEE transactions on magnetics* **2003,** 39, (2), 710-715.
7. Ziegler, M.; Riggert, C.; Hansen, M.; Bartsch, T.; Kohlstedt, H. Memristive Hebbian plasticity model: Device requirements for the emulation of Hebbian plasticity based on memristive devices. *IEEE transactions on biomedical circuits and systems* **2015,** 9, (2), 197-206.
8. Hansen, M.; Zahari, F.; Ziegler, M.; Kohlstedt, H. Double-barrier memristive devices for unsupervised learning and pattern recognition. *Frontiers in neuroscience* **2017,** 11, 91.