Diffusion Models - Applications
04/08/2026

The breadth of scientific applications — from drug discovery to atmospheric modeling, from protein design to seismic inversion — demonstrates that the diffusion paradigm is not merely an image synthesis technique but a general framework for probabilistic inverse problems in science and engineering: wherever one seeks to sample from a complex posterior distribution conditioned on observed data, diffusion models offer a principled and scalable solution.

Drug Discovery

Diffusion models have been adapted from image synthesis to the generation of molecular structures by defining the forward process on three-dimensional atomic coordinates and bond types.

DiffSBDD (*Schneuing et al.*, 2022) and DiffDock (*Corso* et al., 2022) apply diffusion over the SE(3) group — the special Euclidean group of rigid-body motions in 3D space — to generate ligand poses docked into protein binding sites. The forward process adds translational and rotational noise:

$$q(\mathbf{r}_t \mid \mathbf{r}_{t-1}) = \mathcal{N}(\mathbf{r}_t;\ \sqrt{1-\beta_t}\,\mathbf{r}_{t-1},\ \beta_t\mathbf{I}), \quad q(\mathbf{R}_t \mid \mathbf{R}_{t-1}) = \mathrm{IGSO}(3)(\mathbf{R}_{t-1},\ \sigma_t^2)$$

where $\mathbf{r}$ denotes position (translation) and $\mathbf{R}$ denotes orientation (rotation), with the Isotropic Gaussian distribution on SO(3) used for rotational noise. Applications include de novo drug design, generating candidate molecules with desired binding affinity and pharmacological properties — a domain where experimental screening of 10^{60} possible drug-like molecules is physically infeasible.

Weather Forecasting and Climate Emulation

GenCast (*Price et al.*, 2023, Google DeepMind) and CorrDiff (*Mardani et al.*, 2023, NVIDIA) apply diffusion models to ensemble weather forecasting — generating probabilistic distributions over future atmospheric states that capture forecast uncertainty.

Traditional numerical weather prediction (NWP) models solve the *Navier–Stokes* equations discretized on a global grid, requiring supercomputer resources. Diffusion-based emulators treat the atmospheric state (temperature, wind, pressure fields) as images and model the conditional distribution:

$$p_\theta(\mathbf{x}_{t+\Delta t} \mid \mathbf{x}_t)$$

where $\mathbf{x}_t$ is the atmospheric state at time t encoded as a multi-channel field on a global grid. GenCast achieves ensemble forecast skill exceeding ECMWF's (European Centre for Medium-Range Weather Forecasts) ENS system at 15-day horizons, at a fraction of the computational cost.

Medical Imaging

Diffusion models address critical challenges in medical imaging, including data scarcity, privacy constraints on real patient data, and artifact removal.

Reconstruction from undersampled MRI (*Song et al.*, 2021): MRI acquisition time is proportional to the number of k-space measurements. Diffusion models enable compressed sensing reconstruction: given undersampled *Fourier* measurements $\mathbf{y} = \mathcal{A}\mathbf{x} + \boldsymbol{\eta}$ (where $\mathcal{A}$ is a partial *Fourier* operator), the posterior $p(\mathbf{x} \mid \mathbf{y})$ is approximated by conditioning the reverse diffusion process on the data consistency term:

$$\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t \mid \mathbf{y}) = \nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t) + \nabla_{\mathbf{x}_t} \log p(\mathbf{y} \mid \mathbf{x}_t)$$

This enables 4–8× acceleration of MRI scans with diagnostic-quality reconstructions, directly translating to reduced patient burden and scanner throughput improvement.

Anomaly detection (*Wolleb et al.*, 2022): Diffusion models trained on healthy brain MRI can detect pathological structures by measuring the deviation between a patient scan and its diffusion-based reconstruction — pathological regions that violate the learned healthy distribution appear as high residual areas.

Crystal Structure Prediction

CDVAE (*Xie et al.*, 2022) and DiffCSP (*Jiao et al.*, 2023) apply diffusion to crystal structure prediction — determining the stable three-dimensional arrangement of atoms in a material given its chemical composition. The forward process noises both the fractional atomic coordinates and the lattice parameters (unit cell lengths and angles):

$$\mathcal{L}(\mathbf{a}, \mathbf{b}, \mathbf{c}, \alpha, \beta, \gamma) \longrightarrow \mathcal{L}_T \approx \mathcal{N}(\bar{\mathcal{L}}, \sigma_L^2 \mathbf{I})$$

Applications include predicting novel battery electrode materials, high-temperature superconductors, and catalysts for CO_2 reduction — materials synthesis problems of enormous commercial and environmental relevance.

Protein Structure and Dynamics

RFdiffusion (*Watson et al.*, 2023, Nature) applies diffusion over backbone dihedral angles and Cartesian coordinates in SE(3) to de novo protein design — generating protein backbone structures with desired functional properties. The reverse process generates structures conditioned on a target binding site:

$$p_\theta(\mathbf{x}_{\text{protein}} \mid \mathbf{x}_{\text{target}})$$

Experimental validation demonstrated that designed proteins bound their targets with affinities competitive with naturally evolved antibodies — a landmark result illustrating that diffusion models can discover functional structures in a combinatorially vast design space (20^{100} possible 100-residue proteins).

Seismic Imaging

In geophysics, Full Waveform Inversion (FWI) reconstructs subsurface velocity models $\mathbf{m}$ from observed seismic data $\mathbf{d}$, solving the ill-posed inverse problem:

$$\min_{\mathbf{m}} \|\mathcal{F}(\mathbf{m}) - \mathbf{d}\|^2 + \mathcal{R}(\mathbf{m})$$

where $\mathcal{F}$ is the wave-equation forward operator. Diffusion models serve as learned regularizers $\mathcal{R}(\mathbf{m})$ (*Gao & Alkhalifah*, 2023), encoding realistic prior distributions over subsurface geology and enabling probabilistic uncertainty quantification in reservoir characterization — directly impacting oil and gas production decisions and CO_2 storage site selection.

<u>Autonomous Systems and Robotics</u>

Diffusion Policy (*Chi et al.*, 2023, RSS) applies diffusion models to robot learning from demonstration, treating action sequences as generative targets:

$$p_\theta(\mathbf{a}_{t:t+H} \mid \mathbf{o}_{t-k:t})$$

where $\mathbf{a}$ are robot actions, $\mathbf{o}$ are observations, and H is the prediction horizon. By modeling the multi-modal distribution of expert behavior — humans demonstrating a task may approach it in several qualitatively different ways — diffusion policies significantly outperform regression-based imitation learning on dexterous manipulation tasks.

<u>Limitations and Research Frontiers</u>

The sequential nature of the reverse process remains a bottleneck. While DDIM and DPM-Solver reduce steps to $\mathcal{O}(10\text{–}20)$, each step requires a full U-Net forward pass. Consistency Models (*Song et al.*, 2023) attempt to learn a direct mapping from noise to data in one step by enforcing a self-consistency property:

$$\mathbf{f}_\theta(\mathbf{x}_t, t) = \mathbf{f}_\theta(\mathbf{x}_{t'}, t') \quad \forall\, t, t' \text{ on the same ODE trajectory}$$

Diffusion models empirically exhibit excellent mode coverage — they produce diverse samples across the full data distribution. However, classifier-free guidance, while improving individual sample quality, reduces diversity (a manifestation of the fidelity–diversity trade-off encoded in the Precision-Recall framework).

Text-to-image models exhibit semantic hallucination: failure to accurately render object attributes, spatial relationships, or numerical quantities specified in prompts. The CLIP text encoder's limited capacity to encode compositional semantics is a known failure mode under active investigation.

The standard *Gaussian* diffusion framework applies to continuous data. Extension to discrete data (text, molecular graphs) requires non-Gaussian forward processes. Multinomial diffusion (*Austin et al.*, 2021) and absorbing diffusion processes define forward corruption via categorical noise:

$$q(\mathbf{x}_t \mid \mathbf{x}_0) = \mathrm{Cat}\big(\mathbf{x}_t;\ \mathbf{x}_0 \bar{\mathbf{Q}}_t\big)$$

where $\bar{\mathbf{Q}}_t$ is a cumulative transition matrix over the token vocabulary.