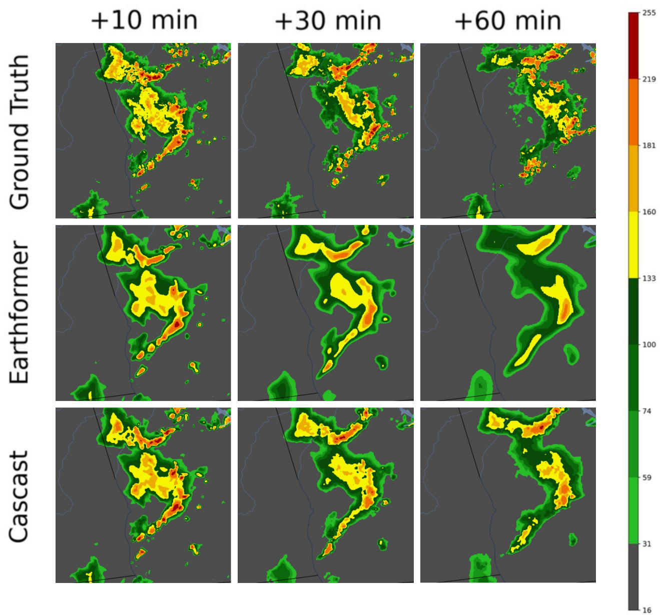
FlowCast: Advancing Precipitation Nowcasting with Conditional Flow Matching

Authors: Bernardo Perrone Ribeiro, Jana Faganeli Pucer

UNIVERZA V LJUBLJANI

The Problem with Precipitation Nowcasting

- Radar-based precipitation nowcasting involves predicting a sequence of future radar images from historical observations
- Deterministic deep learning models optimize for pixel-wise metrics like Mean Squared Error, compelling them to produce a single, best-guess forecast
- As a result, they average over possible futures and produce overly smooth predictions that fail to capture the inherent uncertainty of high-impact events
- While diffusion models offer a sharp, probabilistic alternative, their iterative denoising process is computationally prohibitive for time-critical scenarios

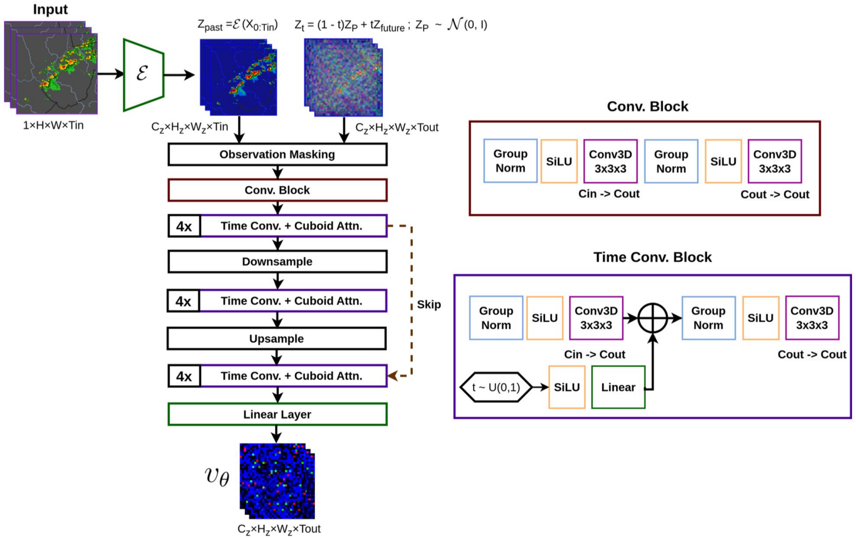


FlowCast: The Core Concept

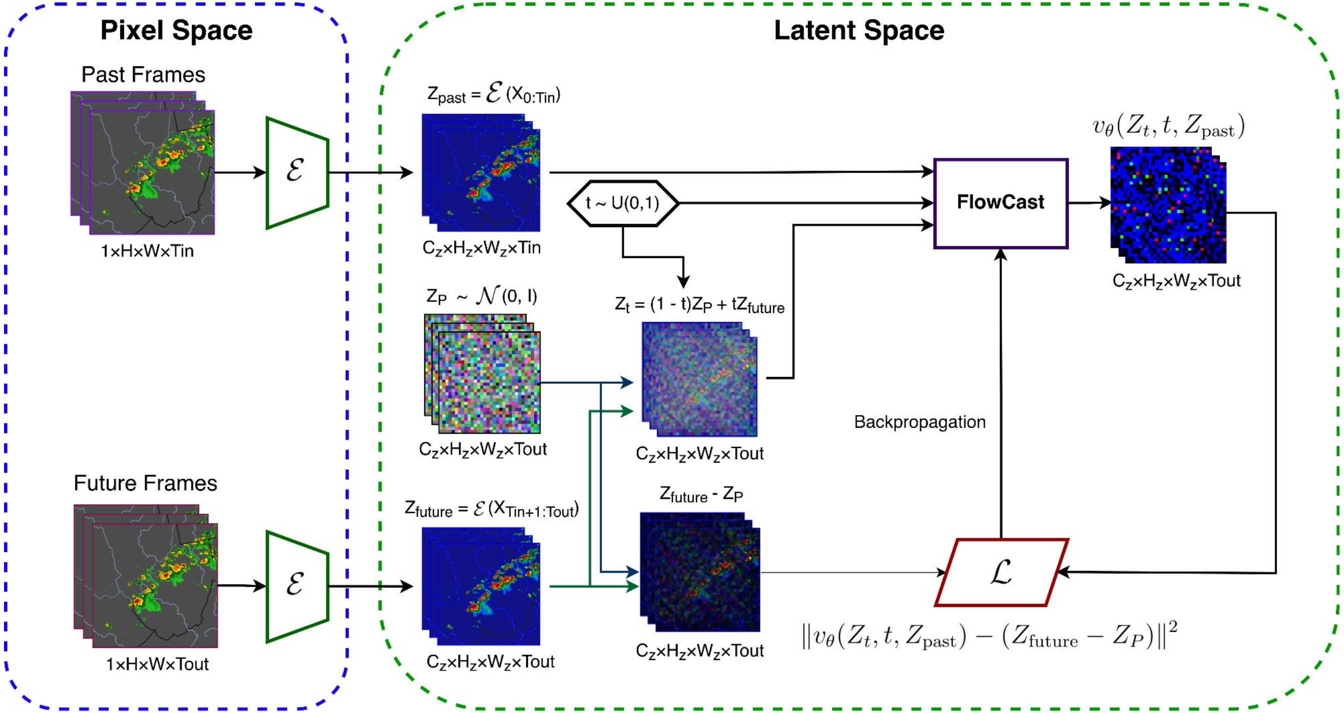
- Standard diffusion models map Gaussian noise to the data distribution via stochastic denoising or curved probability flow ODEs
- This geometric complexity is exactly why diffusion requires hundreds of sampling steps
- To solve this, we introduce FlowCast, the first end-to-end probabilistic model using Conditional Flow Matching (CFM) for nowcasting
- CFM imposes a straight-line ODE prior on the generative process
- This enforces the simplest possible mapping between the noise and data distributions, enabling rapid sampling

The FlowCast Architecture

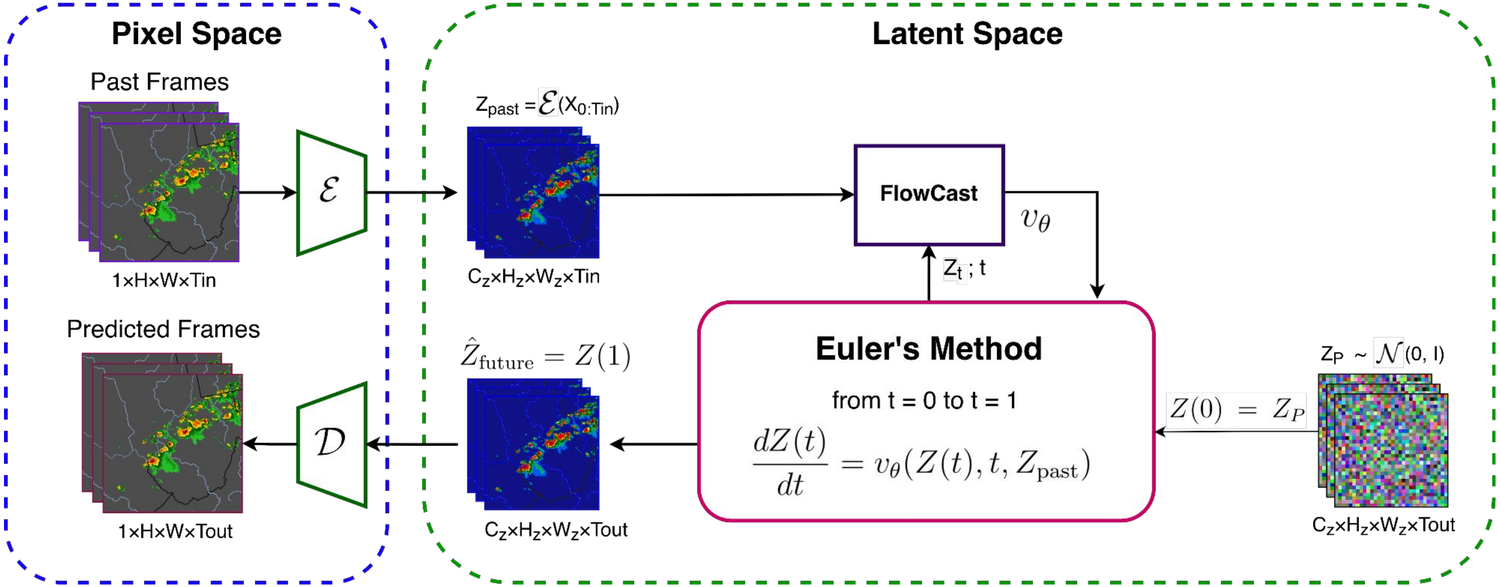
- To efficiently model high-dimensional data, a Variational Autoencoder (VAE) first compresses radar frames into low-dimensional latents
- FlowCast then learns the vector field directly in this compressed latent space
- The model employs a U-Net-like encoder-decoder structure (Earthformer-UNet)
- The core building blocks use Cuboid Attention layers to effectively capture local spatiotemporal dynamics
- By conditioning on the flow time, the model accurately approximates the time-dependent vector field to generate the forecast



FlowCast: Training



FlowCast: Sampling



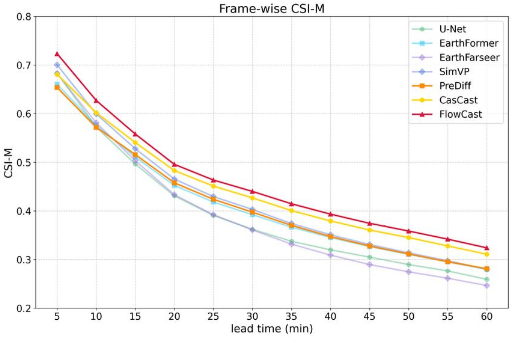
Experiments

- We evaluated FlowCast on the benchmark SEVIR dataset and the local ARSO dataset. Some results for SEVIR:

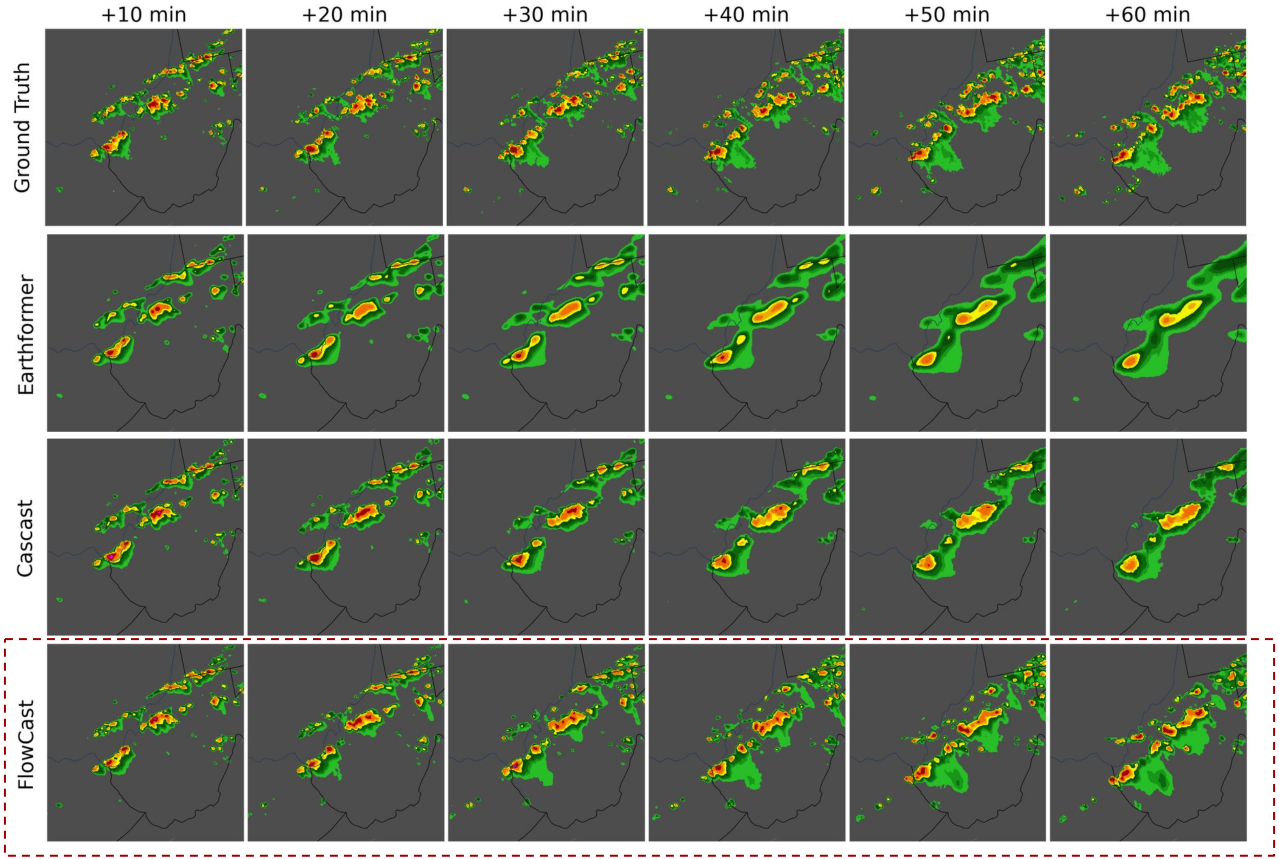
Model	CRPS ↓	Forecast @ 12 steps					Forecast @ +65 min	
		CSI-M ↑	CSI-P16-M ↑	FSS-P16-M ↑	HSS-M ↑	FAR-M ↓	CSI-M ↑	CSI-219 ↑
U-Net	0.0273	0.394	0.384	0.661	0.497	0.308	0.259	0.009
Earthformer	0.0252	0.411	0.407	0.686	0.518	**0.285**	0.280	0.016
Earthfarseer	0.0256	0.389	0.393	0.636	0.486	0.289	0.247	0.001
SimVP	0.0249	0.423	0.424	0.701	0.532	0.298	0.280	0.012
PreDiff	0.0189	0.413	0.423	0.699	0.523	0.313	0.281	0.018
CasCast	0.0201	0.442	**0.520**	0.763	0.562	0.383	0.311	0.054
FlowCast	**0.0182**	**0.460**	0.506	**0.767**	**0.580**	0.325	**0.324**	**0.057**

Extreme Events:

Model	SEVIR					
	CSI		HSS		FAR	
	181	219	181	219	181	219
U-Net	0.205	0.122	0.314	0.193	0.366	0.508
Earthformer	0.229	0.109	0.348	0.180	0.354	**0.343**
Earthfarseer	0.194	0.097	0.291	0.152	**0.341**	0.412
SimVP	0.244	0.137	0.365	0.220	0.370	0.404
PreDiff	0.237	0.128	0.361	0.206	0.384	0.467
CasCast	0.286	0.195	0.427	0.309	0.501	0.567
FlowCast	**0.301**	**0.202**	**0.443**	**0.317**	0.425	0.482



Experiments



Experiments

- A direct ablation study reveals the CFM objective is significantly more accurate and efficient than a diffusion objective on the exact same architecture

Model	CRPS ↓	CSI-M ↑	CSI-P16-M ↑	FSS-M-P16 ↑	HSS-M ↑	FAR-M ↓	Time/Seq. (s)
CFM (1 steps)	0.0207	0.454	0.504	0.763	0.571	0.337	**2.6**
CFM (10 steps)	**0.0168**	**0.455**	**0.514**	**0.764**	**0.572**	0.338	24
DDIM (10 steps)	0.0262	0.395	0.450	0.622	0.503	0.335	24
DDIM (50 steps)	0.0212	0.398	0.451	0.635	0.504	0.321	120
DDIM (100 steps)	0.0208	0.398	0.450	0.664	0.502	**0.319**	239