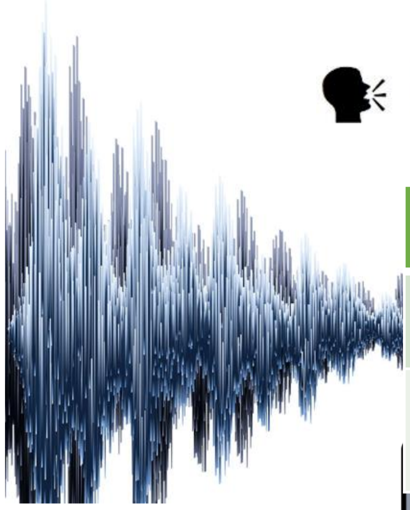
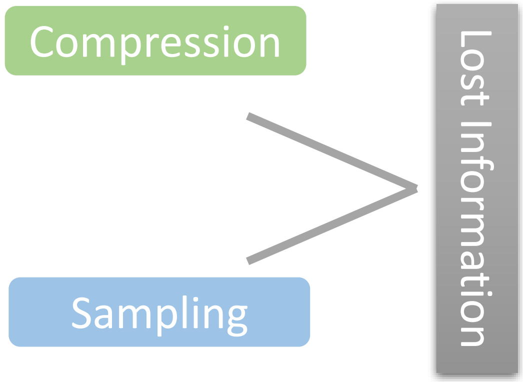
Modeling Lost Information in Signal Processing

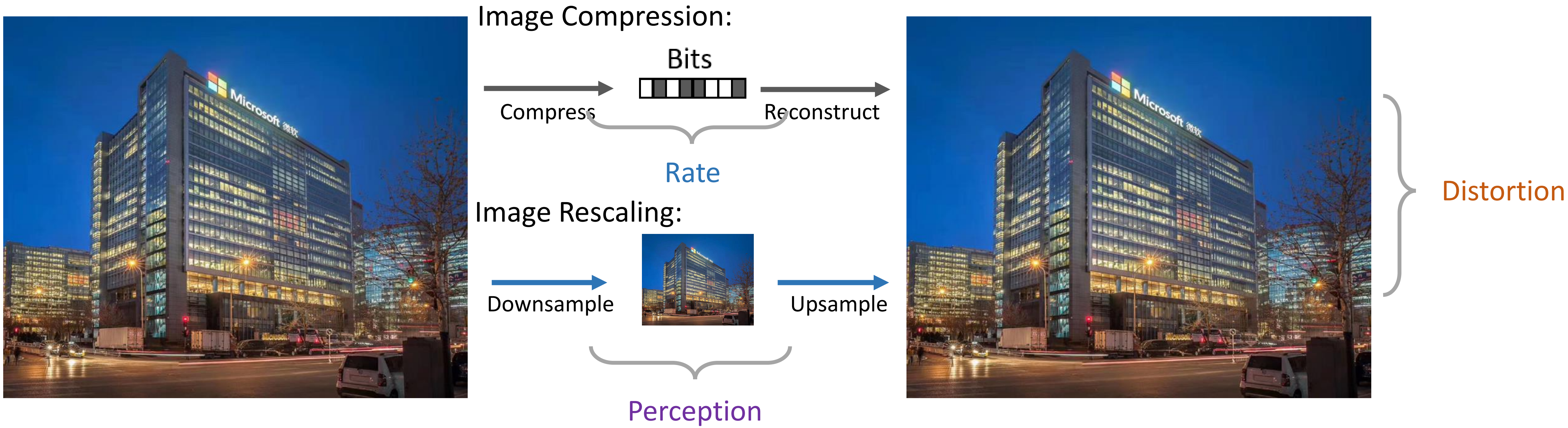
Shuxin Zheng, Machine Learning Group, MSRA

Signal (Processing) is Everywhere

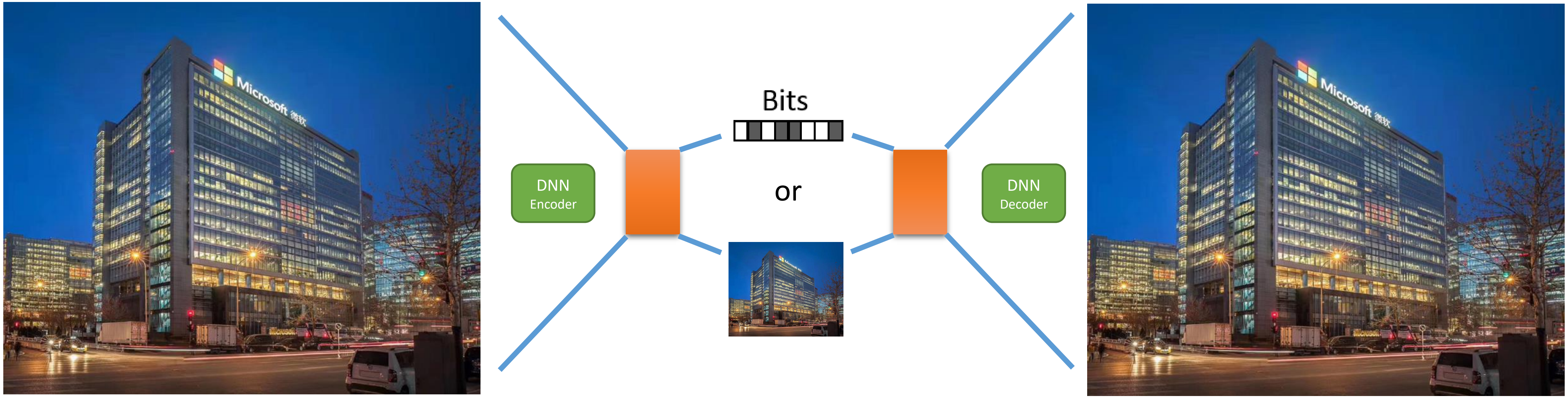


Task	Sameness	Modality	Optimization
Image Compression	Low-dimensional	Bitstream	**Rate-Distortion**
Image Rescaling		Low-resolution	**Perception-Distortion**



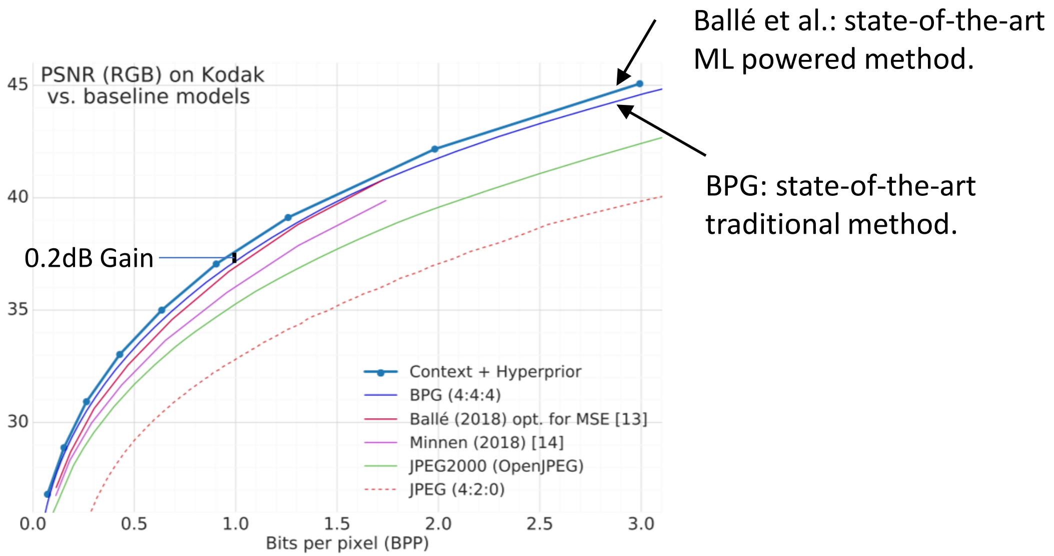


Signal Processing with Machine Learning



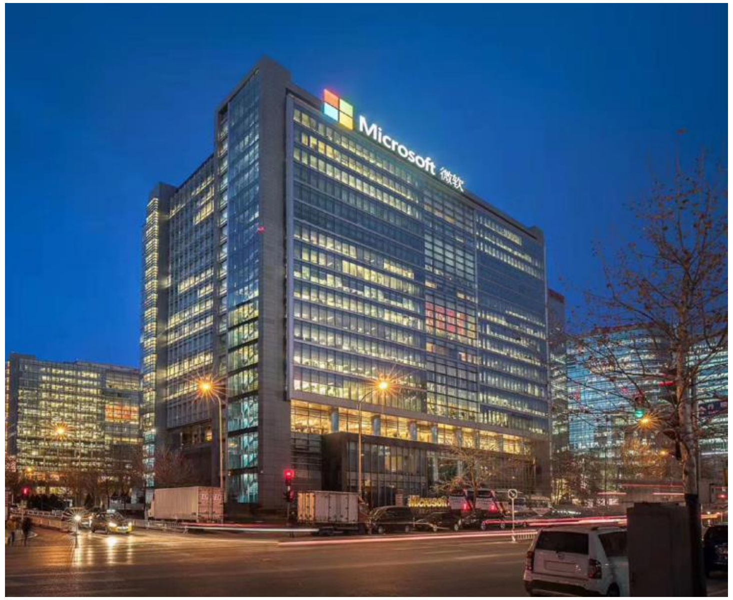
- **ML powered** signal processing becomes a hot topic and has achieved better performance than traditional method.

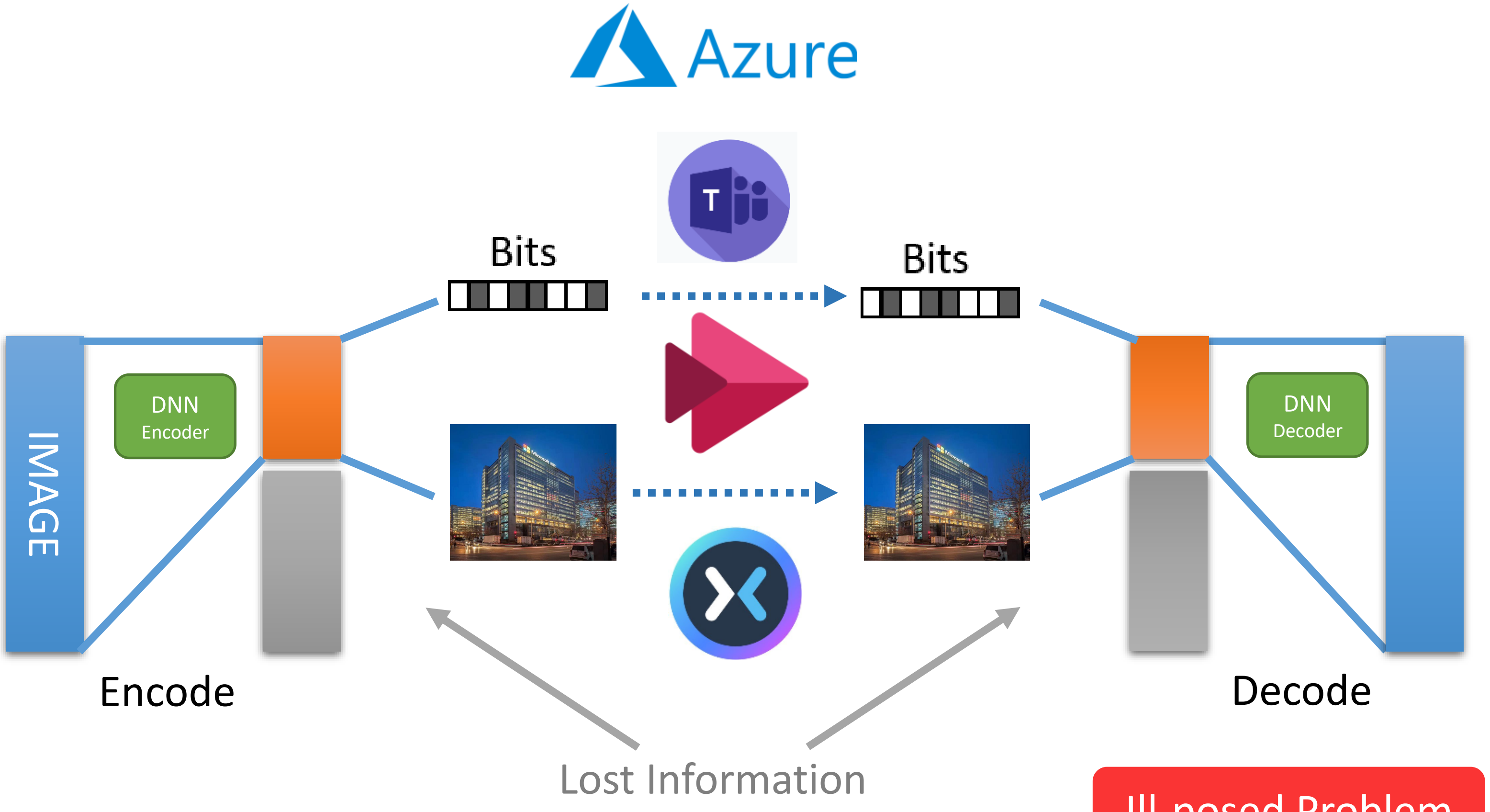
- **Google (Ballé et al.)**[1][2][3] holds the leading position in lossy image compression for years.



[1] **Ballé J**, Laparra V, Simoncelli E. End-to-end optimized image compression[C]//5th International Conference on Learning Representations, ICLR 2017.
[2] **Ballé J**, Minnen D, Singh S, et al. Variational image compression with a scale hyperprior [C]//6th International Conference on Learning Representations, ICLR 2018.
[3] Minnen D, **Ballé J**, Toderici G D. Joint autoregressive and hierarchical priors for learned image compression[C]//Advances in Neural Information Processing Systems. 2018: 10771-10780.

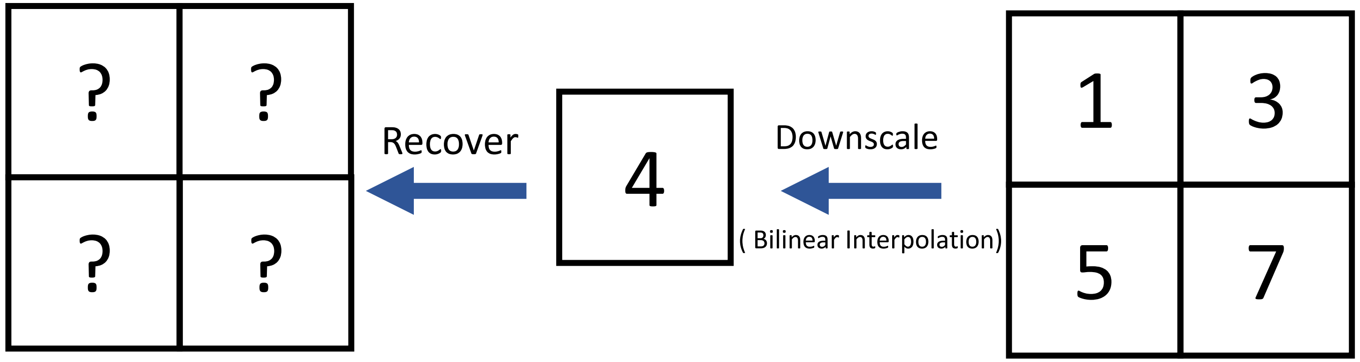
Lost Information



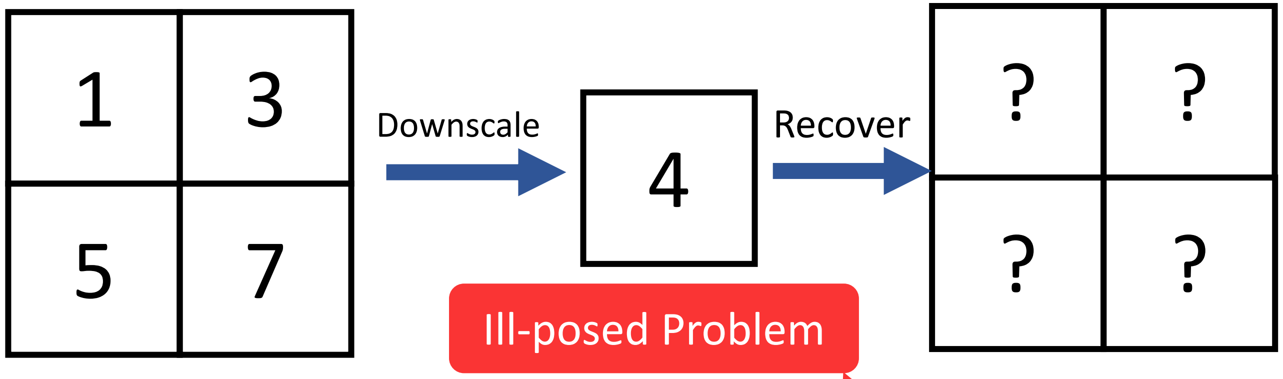


Ill-posed Problem

Lost Information
↓
Ill-posed Problem



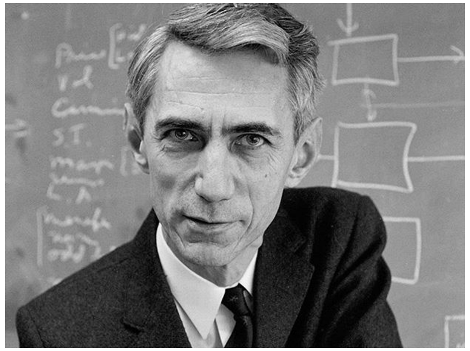
Capture Lost Information



What **information** is **lost** during downsample?

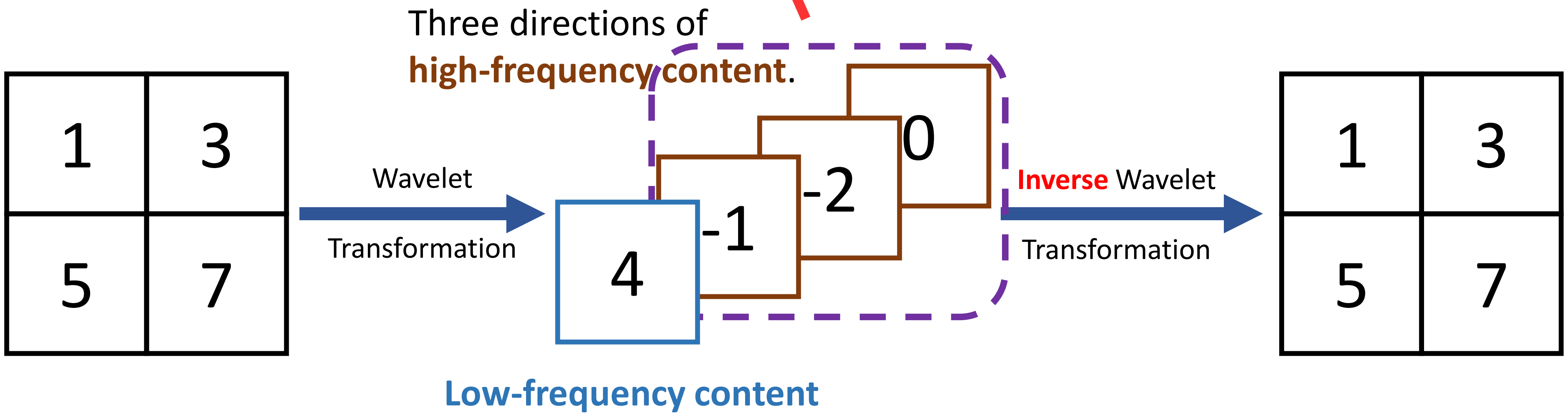
*"**High-frequency content** will get **lost** during sample rate conversion."*

--According to Nyquist-Shannon Sampling Theorem

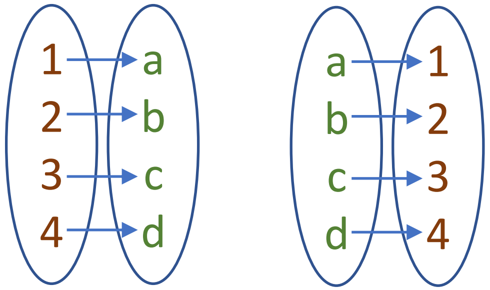


Claude Shannon

Harry Nyquist

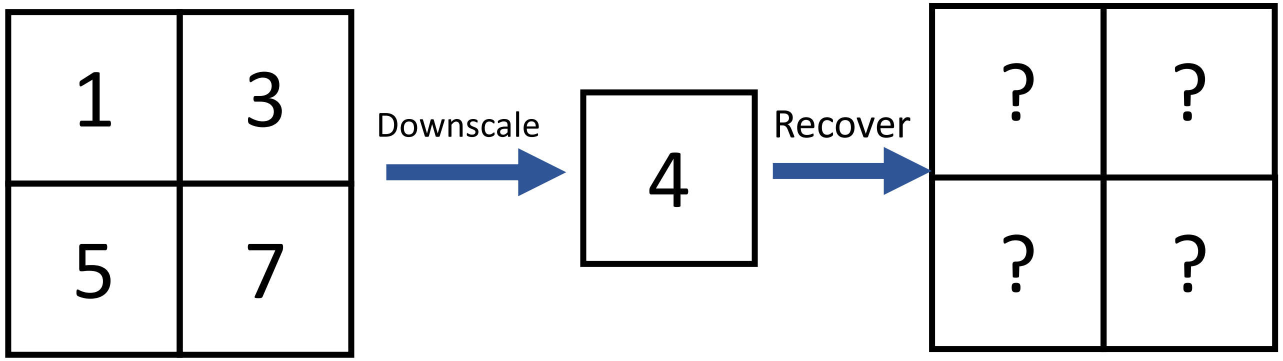


$f(x) = y$ $\quad$ $f^{-1}(y) = x$



Inverse Function

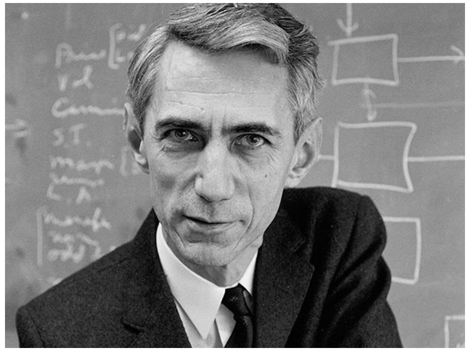
Capture Lost Information



What **information** is **lost** during downsample?

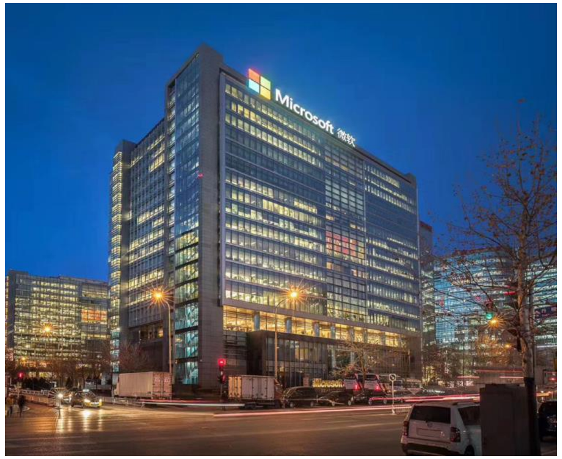
*"**High-frequency content** will get **lost** during sample rate conversion."*

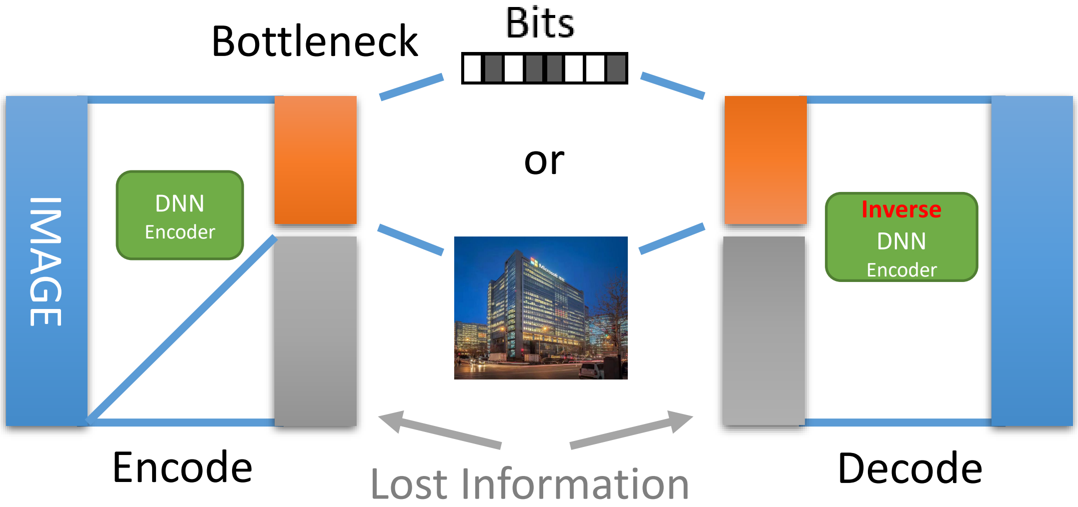
--According to Nyquist-Shannon Sampling Theorem


Claude Shannon

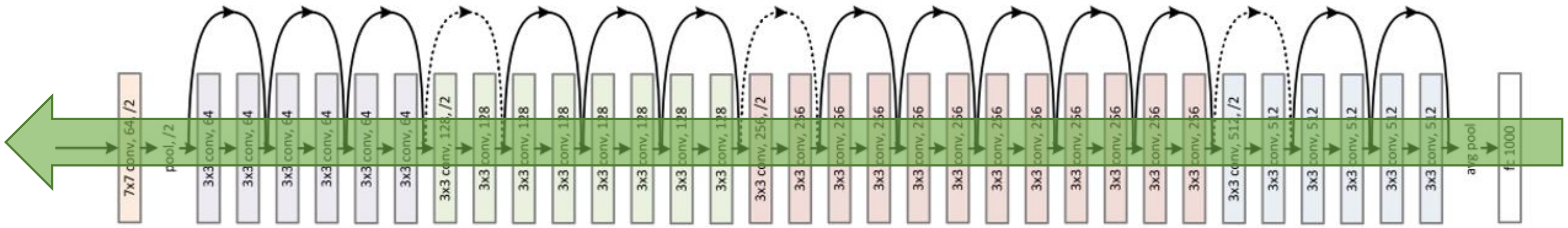

Harry Nyquist

What is the **inverse function** of a **DNN** model?

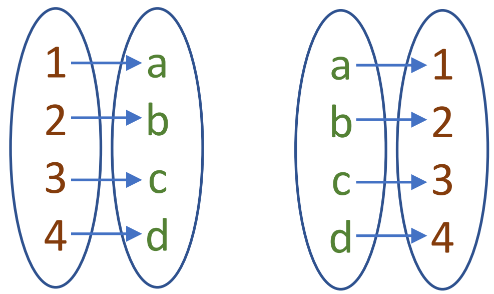




Quick Recap of Invertible Neural Network

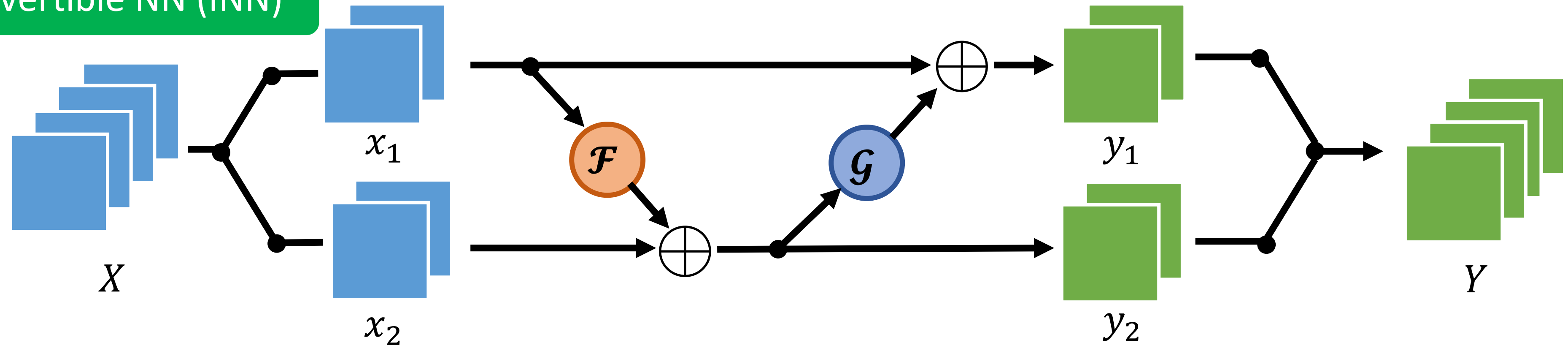


$f(x) = y$ $\qquad f^{-1}(y) = x$



Inverse Function

Invertible NN (INN)



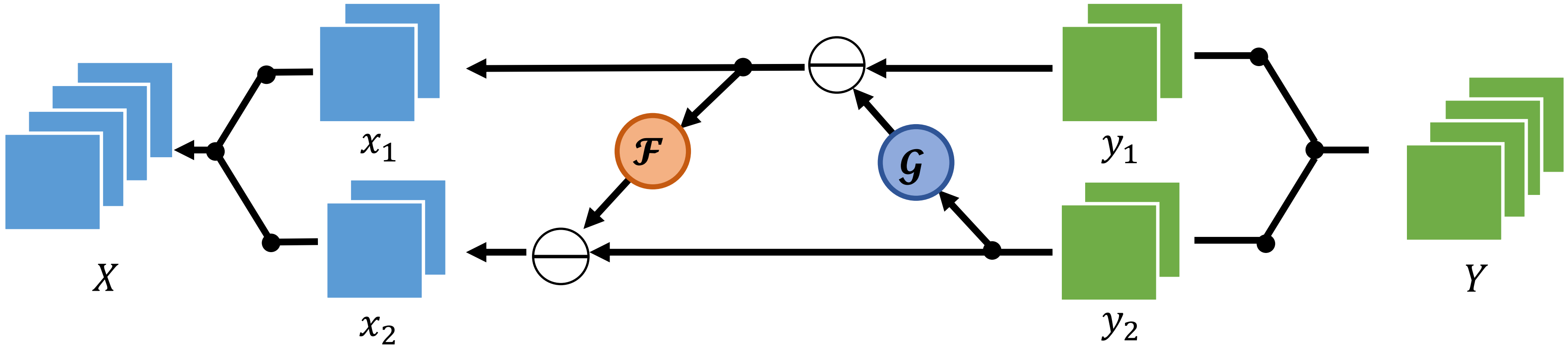
Forward Pass

$$y_1 = x_1 + \mathcal{F}(x_2)$$
$$y_2 = x_2 + \mathcal{G}(y_1)$$

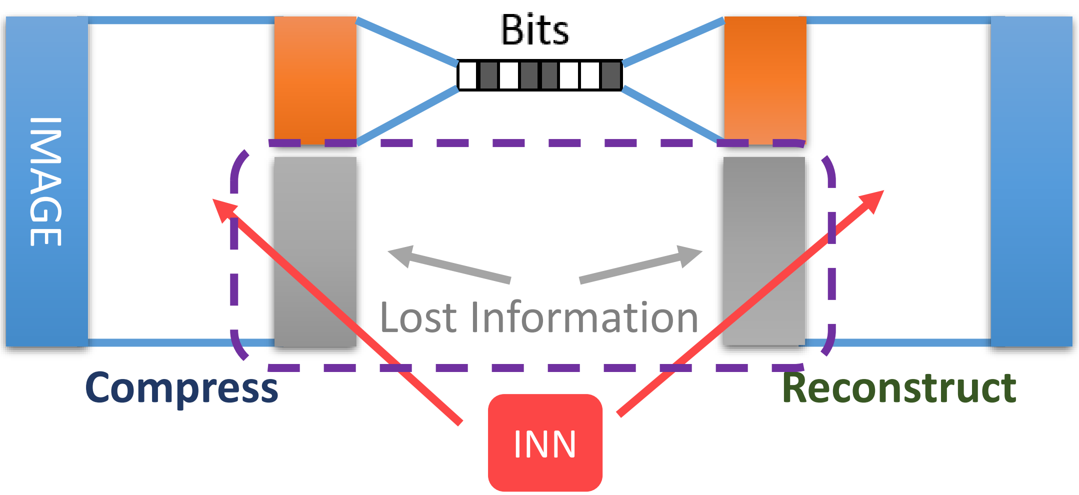
Strictly Invertible

Inverse Pass

$$x_2 = y_2 - \mathcal{G}(y_1)$$
$$x_1 = y_1 - \mathcal{F}(x_2)$$



Face the "Lost Information" Challenge



Question:

If we (_____) the lost information…

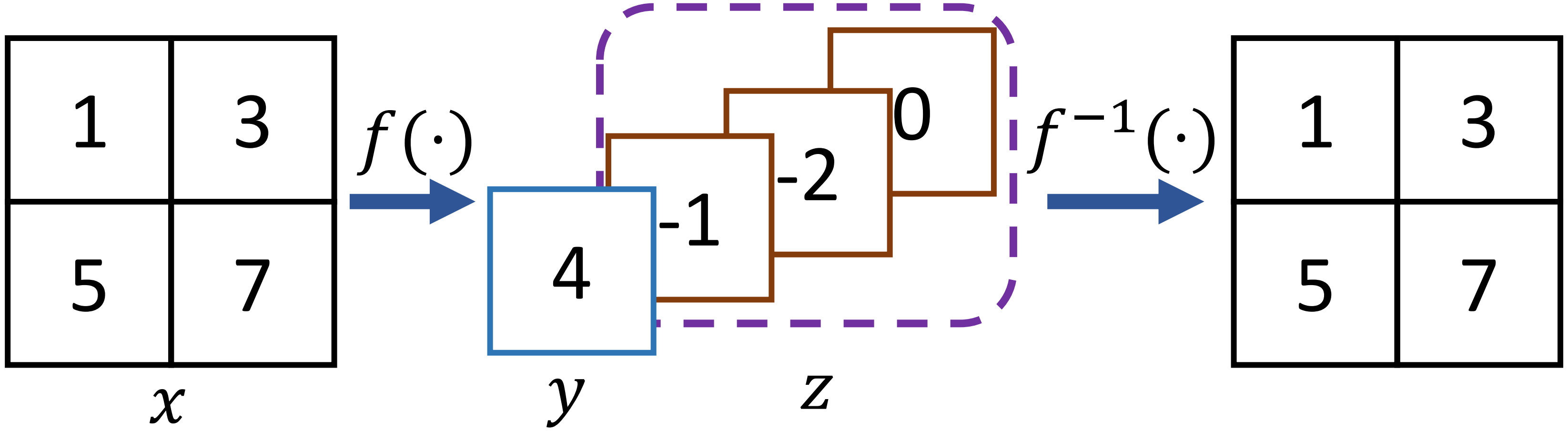
(A). Preserve

How to **compress**?

(B). Abandon

How to **reconstruct**?

Do we have a better choice?

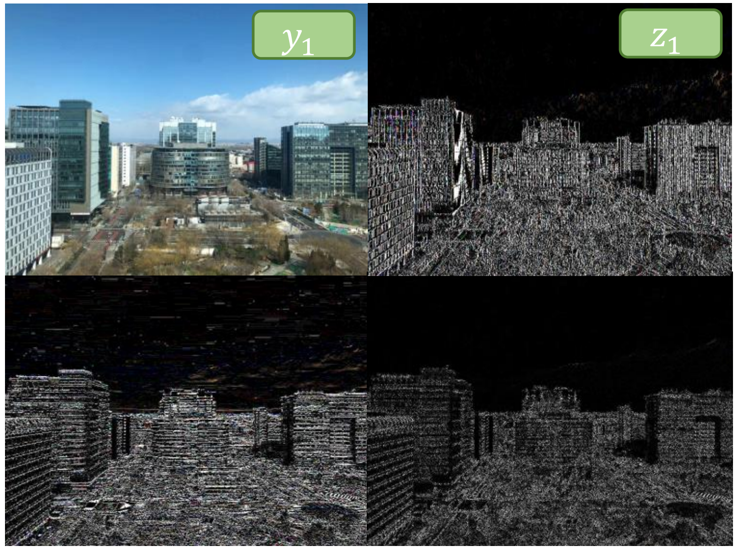


Why z can't be abandoned?

$$f(p(x)) = p(y)p(z|y)$$

Answer: z is case-specific!

Model the Case-Specific Lost Information



$$f\big(p(x)\big) = p(y)p(z|y)$$

Reparameterization

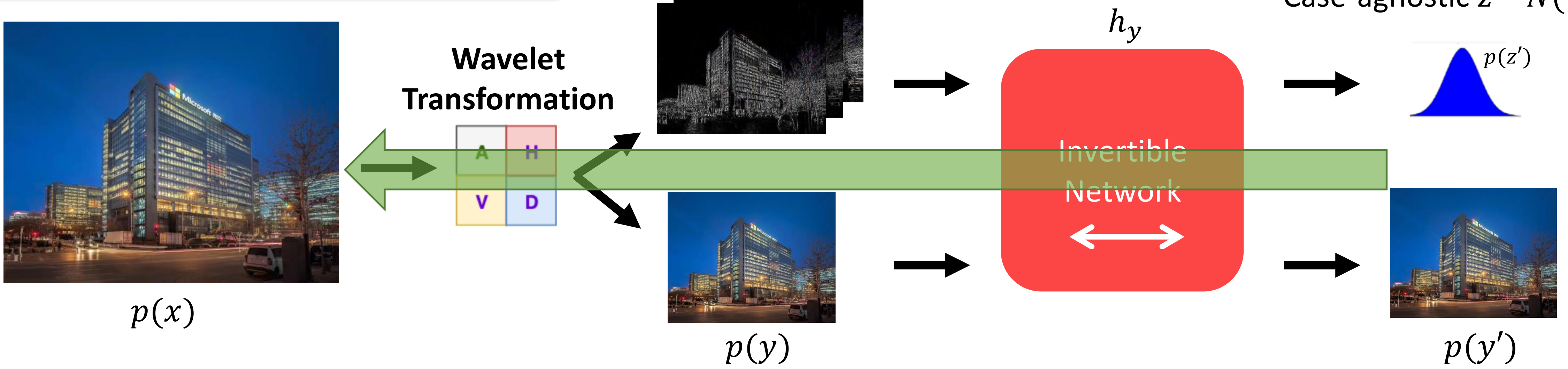
Let $z' = h_y(z)$,
when $z \sim p(z|y)$,
$z' \sim N(0, I_k)$.

Case-agnostic:
Follow a Gaussian!
Independent from y!

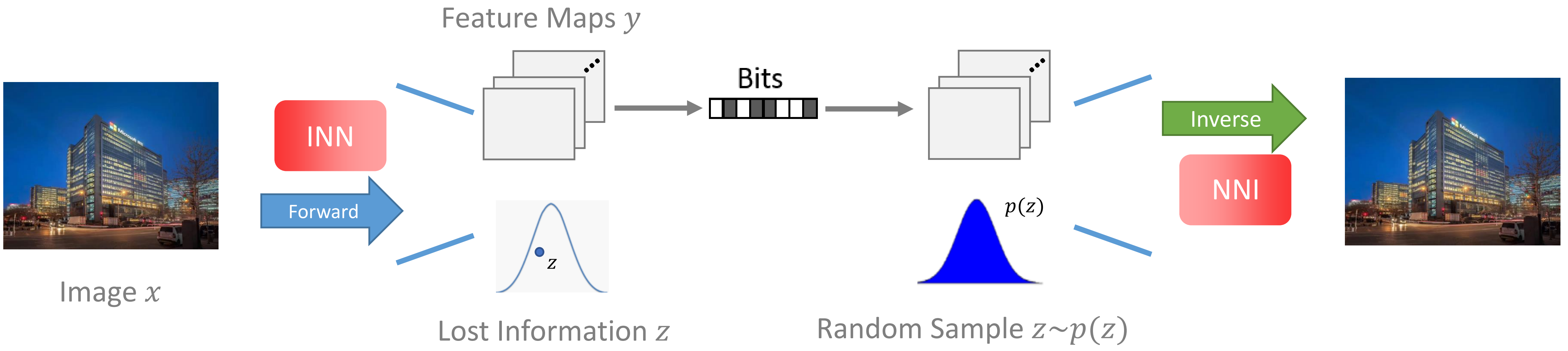
Safely abandon z, before sending y
Easily find back z, after receiving y

Case-specific $z \sim p(z|y)$

Case-agnostic $z' \sim N(0, I_k)$



Modeling Lost Information

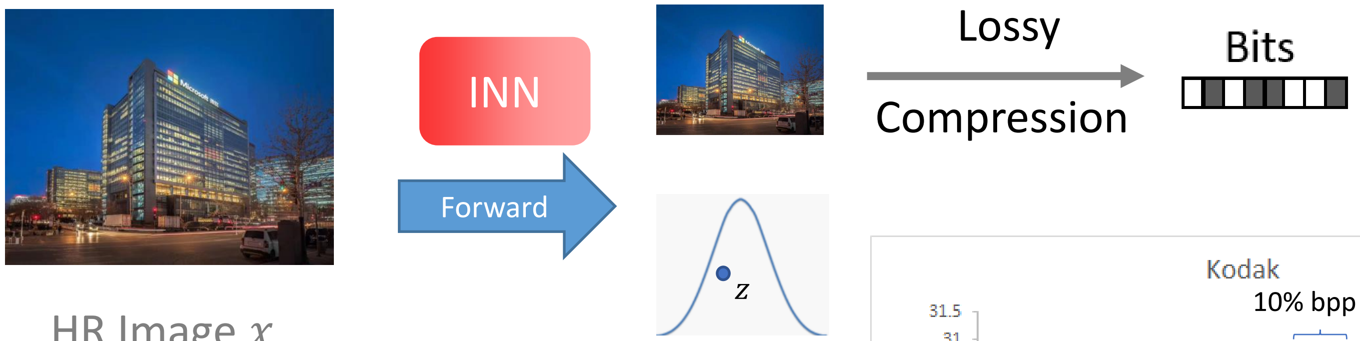


Training Objective	Rescaling	Compression
Distortion	L_1 or L_2 Reconstruction Loss	
Perception	LR Guidance	
Rate		Bitrate (likelihood(y))
Distribution Matching	Backward JS Divergence Loss	

Rescaling Performance (PSNR)

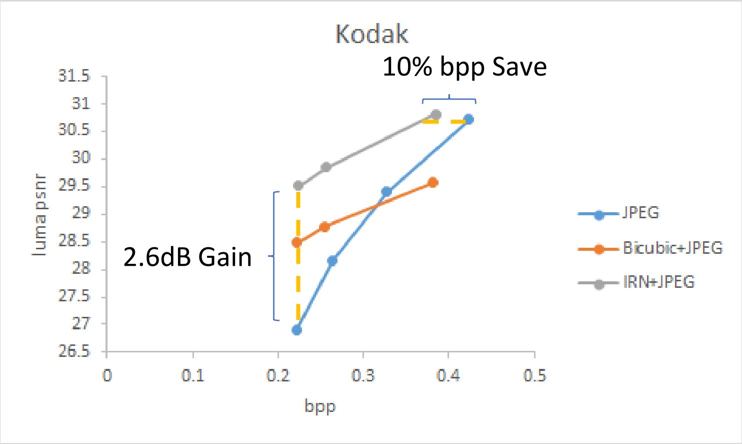
PSNR (dB)	**Scale**	**Param↓**	**Set5↑**	**Set14↑**	**BSD100↑**	**Urban100↑**	**DIV2K val↑**
Bicubic & RCAN ECCV 2018	2x	15.4M	38.27	34.12	32.41	33.34	-
Bicubic & SAN CVPR 2019	2x	15.7M	38.31	34.07	32.42	33.10	-
TAD & TAU ECCV 2018	2x	-	37.69	33.90	32.62	31.96	36.13
CAR & EDSR **(SOTA)** TIP 20	2x	51.1M	38.94	35.61	33.83	35.24	38.26
Ours MSRA, ECCV20	2x	1.66M	43.99 (+5.05dB)	40.79 (+5.18dB)	41.32 (+7.49dB)	39.92 (+4.68dB)	44.32 (+6.06dB)
Bicubic & RCAN ECCV 2018	4x	15.6M	32.63	29.0	27.84	27.03	30.92
Bicubic & SAN CVPR 2019	4x	15.7M	32.64	28.92	27.78	26.79	-
TAD & TAU ECCV 2018	4x	-	31.59	28.36	27.57	25.56	30.25
CAR & EDSR **(SOTA)** TIP 20	4x	52.8M	33.88	30.31	29.15	29.28	32.82
Ours MSRA, ECCV20	4x	4.35M	36.19 (+2.31dB)	32.67 (+2.36dB)	31.64 (+2.49dB)	31.41 (+2.13dB)	35.07 (+2.25dB)

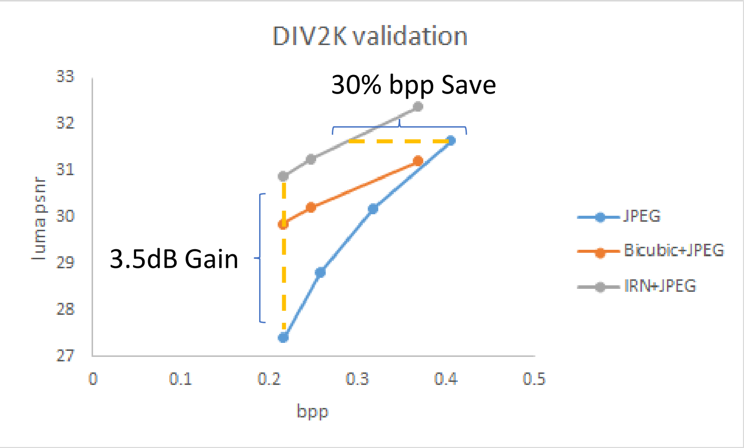
Compression Performance (PSNR/bpp)



HR Image x

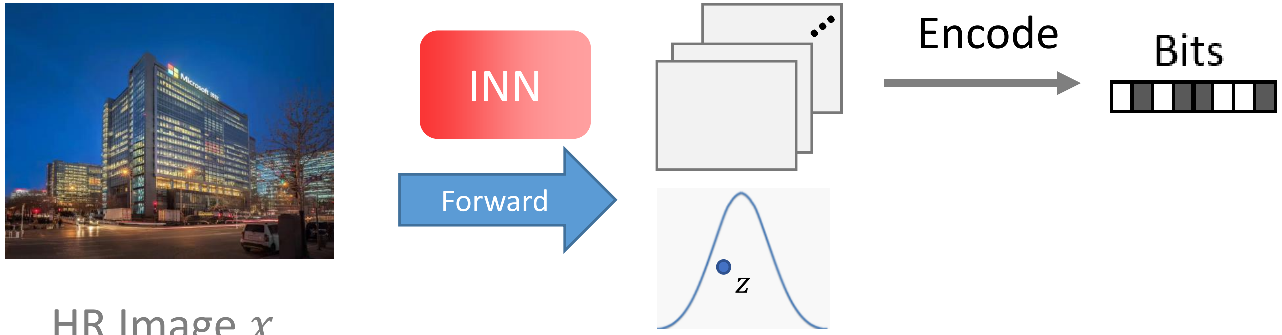
When Compress a High-Resolution Image into Low Bitrate:
downscale + compress > compress



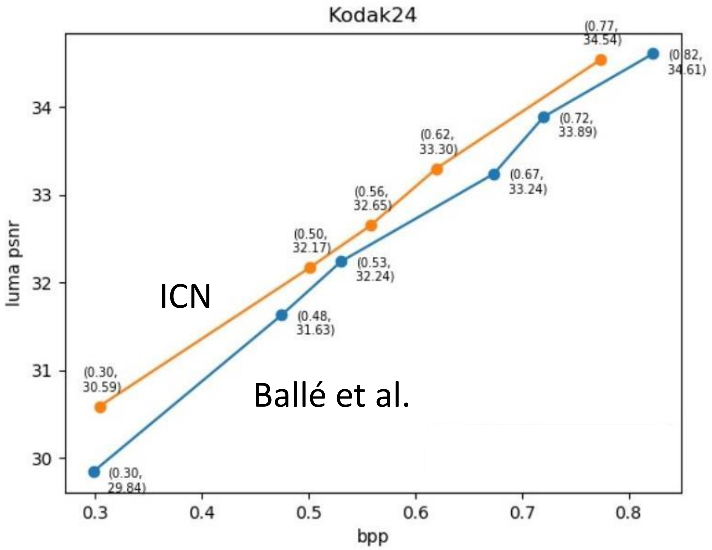


IRN: Invertible Rescaling Net

ICN: Invertible Compression Net



HR Image x



Thanks!

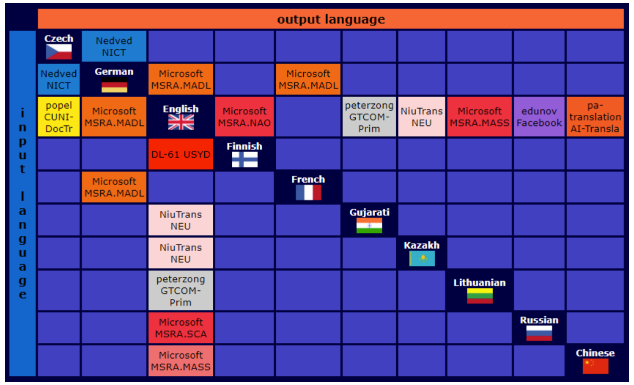
The code is available: https://github.com/pkuxmq/Invertible-Image-Rescaling
For any further question, contact: **shuz**@microsoft.com

Machine Learning Group

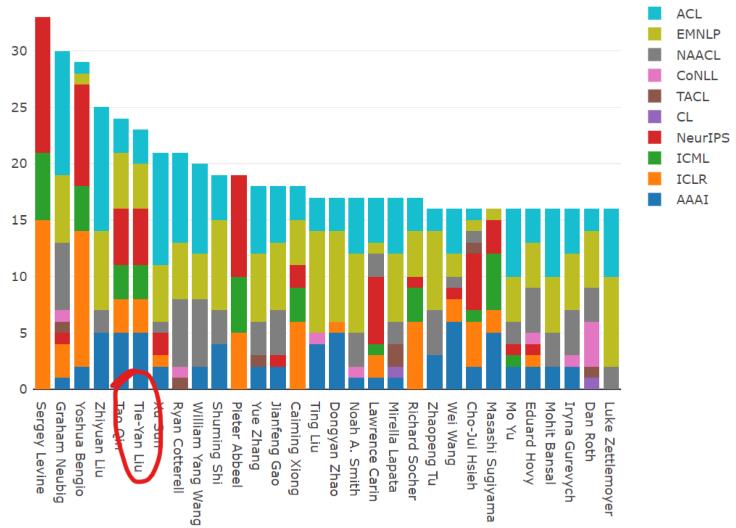
Lead by Prof. Tie-Yan Liu, focus on fundamental and innovative machine learning research, including machine learning theory, algorithms and applications. Actively contribute to academic community. Conduct many impactful techniques into Microsoft products, including Bing, Advertising, Xbox, Azure etc.

SUPHX

Super-human Mahjong AI



First place in WMT 2019



High Research Impact