



# Creating New Realities: Possibilities, Limitations and Impact of Generative Image AI

**Philipp Bauer**  
Sr. Software Architect

# About Philipp



Sr. Software Architect @ CODE Consulting

15+ Years of Experience

Certified Media Designer DE

[pbauer@codemag.com](mailto:pbauer@codemag.com)



# CODE30

YEARS



30 years of  
***"Helping People Build Better Software"***

# About CODE Consulting



Custom software development  
(web, mobile, desktop and cloud platform apps)

Copilot development, AI, GPT, Azure OpenAI,  
machine learning, and more...

Project rescue (cloud, web, mobile, Windows, ...)

App modernization (primarily Visual FoxPro)

Support & maintenance for existing applications



# AI Consulting Services



Check out our new Executive Briefing offer!

We can help with your AI needs

What does AI mean for you?

“Skunk Works” Projects

[codemag.com/AI](https://codemag.com/AI)

[codemag.com/ExecutiveBriefing](https://codemag.com/ExecutiveBriefing)



# CODE Staffing



Disrupting the world of staffing!

Giving our customers the ability to have staff on par with Silicon Valley companies ...

... and our employees a work environment in a bleeding-edge tech company with the **industry leading benefits!**

[codestaffing.com](http://codestaffing.com)



# In today's Session



Capabilities and limitations of image generators

How do latent diffusion models work?

Impact of generative AI

Conclusion

A large, fluffy white rabbit is sitting in the middle of a city street, facing the camera. The rabbit is positioned in the center of the frame, with its front paws resting on its lap. The background shows a busy city street with several cars, including a red Honda Civic on the left and a silver SUV on the right. Tall buildings line the street, and a traffic light is visible on the right. The overall scene is slightly dimmed, with the text overlaid in the center.

# Capabilities and Limitations



# Realism

CODE  
PRESENTS



Photo of a burger with cheese and fries, food photography, foreground focus, epicurious, volumetric light, background is blurry restaurant, warm lighting



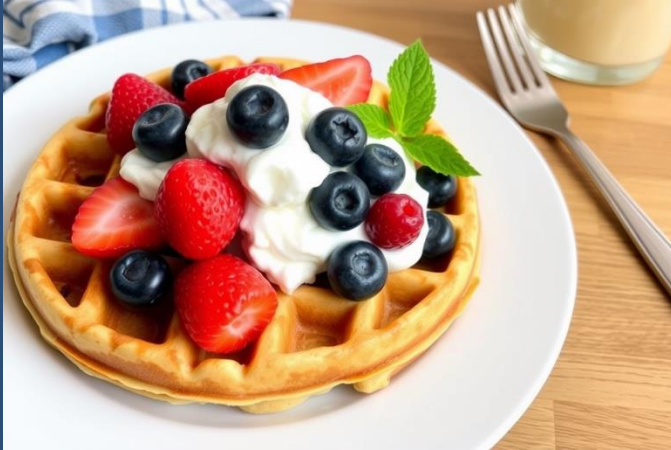
close up photograph of sushi on wooden cutting board, hands level, presenting hand gesture, in the background a blurry apron



A colorful image of vegetable fried rice with a side of hard-boiled eggs and herbs. The rice is garnished with chopped vegetables like zucchini, bell peppers, and onions, all cooked together. The eggs have been boiled until the yolks are solid yellow ...

# Realism

**CODE**  
PRESENTS





# Realism?

**CODE**  
PRESENTS





# Trainability

**CODE**  
PRESENTS

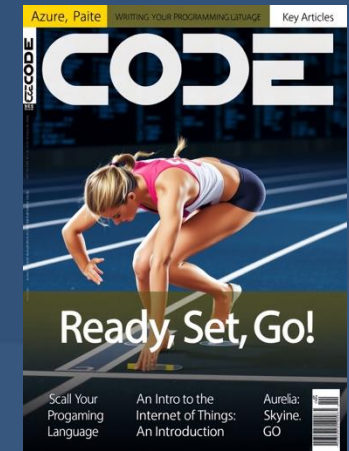
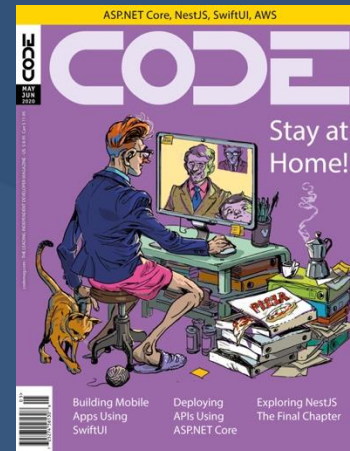
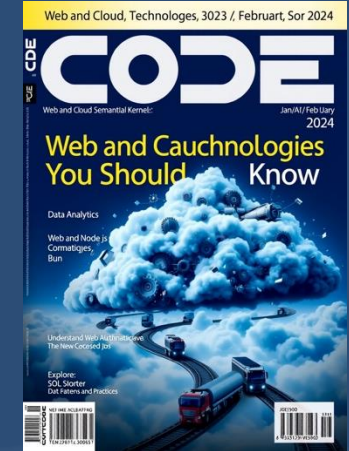
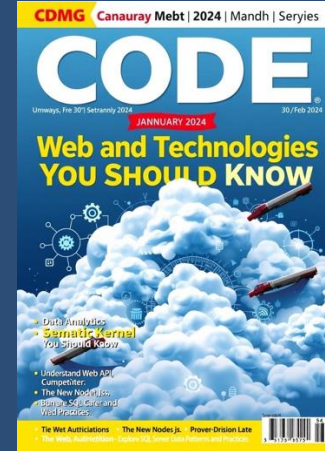
Original Output



Training Image



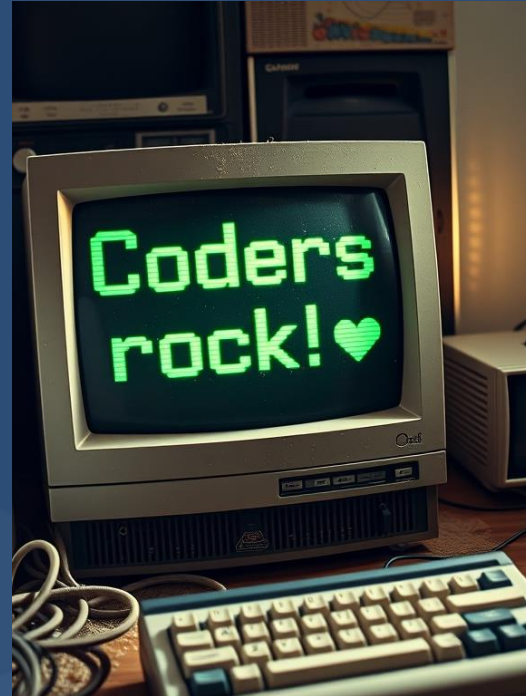
LoRA Output





# Text Handling

**CODE**  
PRESENTS



# Text Handling

**CODE**  
PRESENTS





# Colorization

**CODE**  
PRESENTS





# Upscaling

**CODE**  
PRESENTS





# Animations / Video



We orbit around lighthouse, hyper dynamic movement in orbiting motion, shallow depth of field, subject in focus. In a dark and gloomy style

runway

# Prompt Adherence

**CODE**  
PRESENTS



This is an image of a magazine cover. The cover features the title "CODE" in large, bold letters at the top, with additional text below that reads "Web and Cloud Technologies You Should Know." The number 30 is prominently displayed on the cover, indicating the issue or edition number.

The background of the cover is a surreal, digital-art style image depicting a cityscape with futuristic elements. There are floating trains, clouds, and what appear to be digital representations of various objects like buildings, clocks, and other abstract shapes. The overall color scheme is predominantly cool blues and whites, giving the artwork an ethereal, high-tech feel.

At the top of the cover, there's a mention of "Data Analytics, Semantic Kernel, Bun, .NET Core," which could be related to the magazine's content, or specific topics covered inside.

The date "Jan 2024" is also present on the

cover, suggesting when this particular issue was released. The logo of a train in the background could symbolize connectivity or transportation technologies.

The bottom left corner of the image includes the barcode for scanning and digital authentication, along with the text "Understand Web API Authentication." This suggests that one of the articles in the magazine may be about web application programming interface (API) authentication methods.

Lastly, there's a mention of "The New Node.js Explorer," which could indicate an article or feature related to Node.js, a popular JavaScript runtime built on Chrome's V8 JavaScript engine.

# Prompt Adherence

**CODE**  
PRESENTS



Create a CDMG magazine cover for "CODE" with a focus on web and cloud technologies.

The central design features large, futuristic clouds representing cloud computing, filled with icons like gears, databases, and network nodes. Trains and trucks move through the clouds, symbolizing data infrastructure.

Use cool tones like blues and greys for a modern, tech-forward feel. Include text elements like "Web and Cloud Technologies You Should Know" as the main headline, and a yellow banner at the top with "Data Analytics, Semantic Kernel, Bun, .NET Core."

Add smaller headlines like "Understand Web API Authentication" and "Explore SQL Server Data Patterns." Include "30 Years" and "Jan/Feb 2024" on the cover.



# LLM-Enhanced Prompts

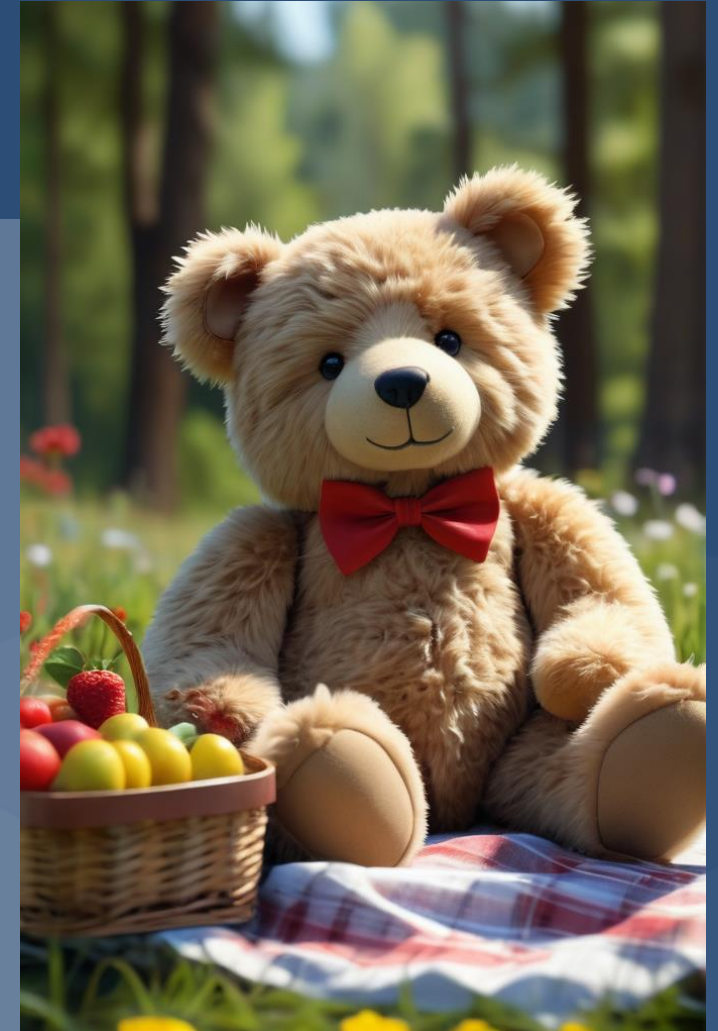
CODE  
PRESENTS



teddy bear

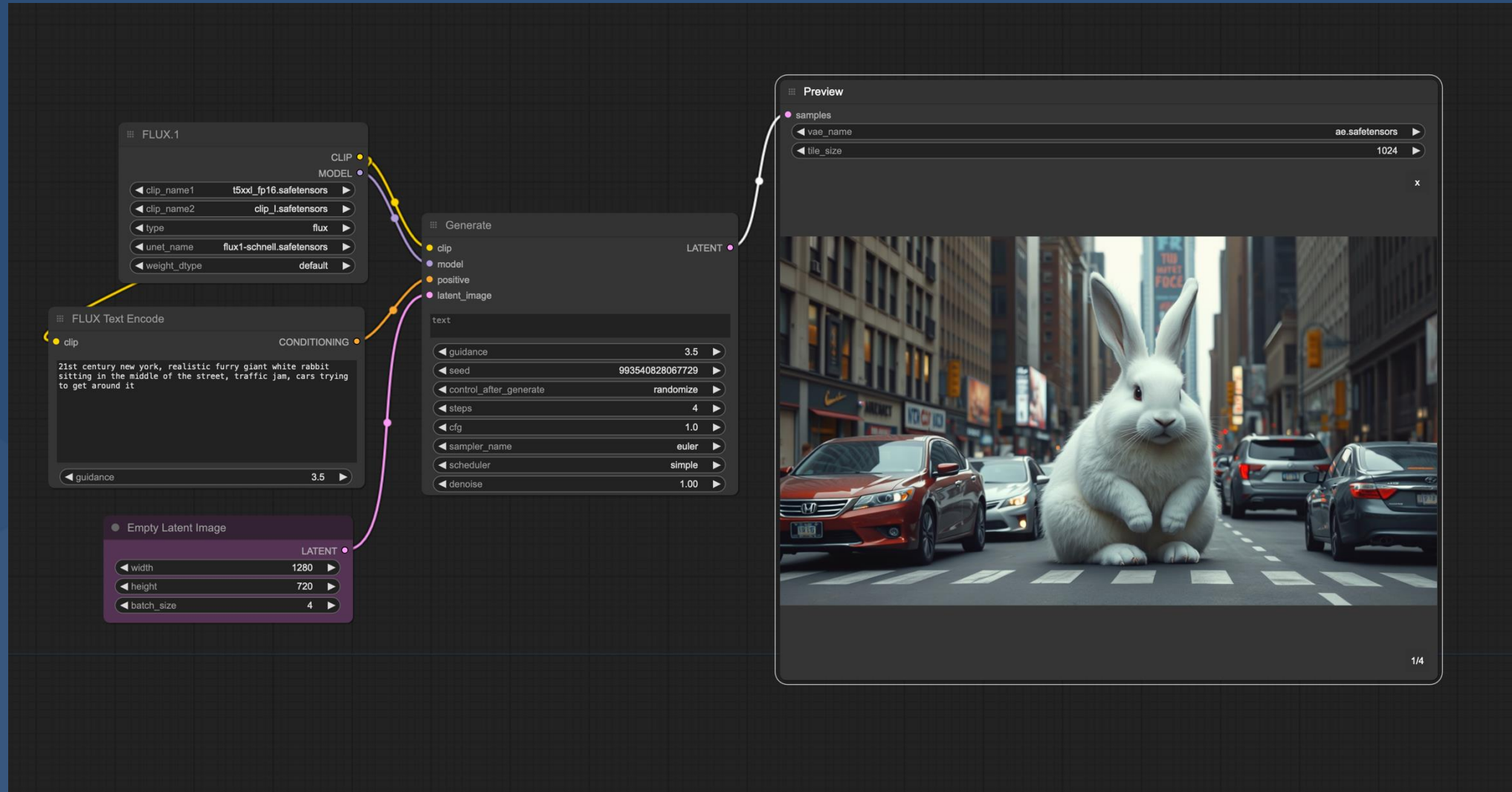
enhance!

A soft, cuddly, and adorable teddy bear with expressive button eyes, sitting in a sunlit meadow, wearing a red bow tie, holding a small picnic basket filled with colorful fruits. The teddy bear's fur is fluffy and smooth to touch, with every strand visible in high definition detail. The background is a serene forest landscape with flowers blooming all around.





# Comfy UI



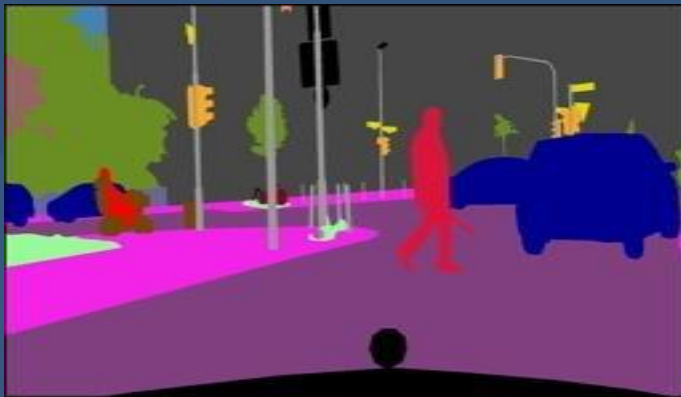


# How do Latent Diffusion Models work?

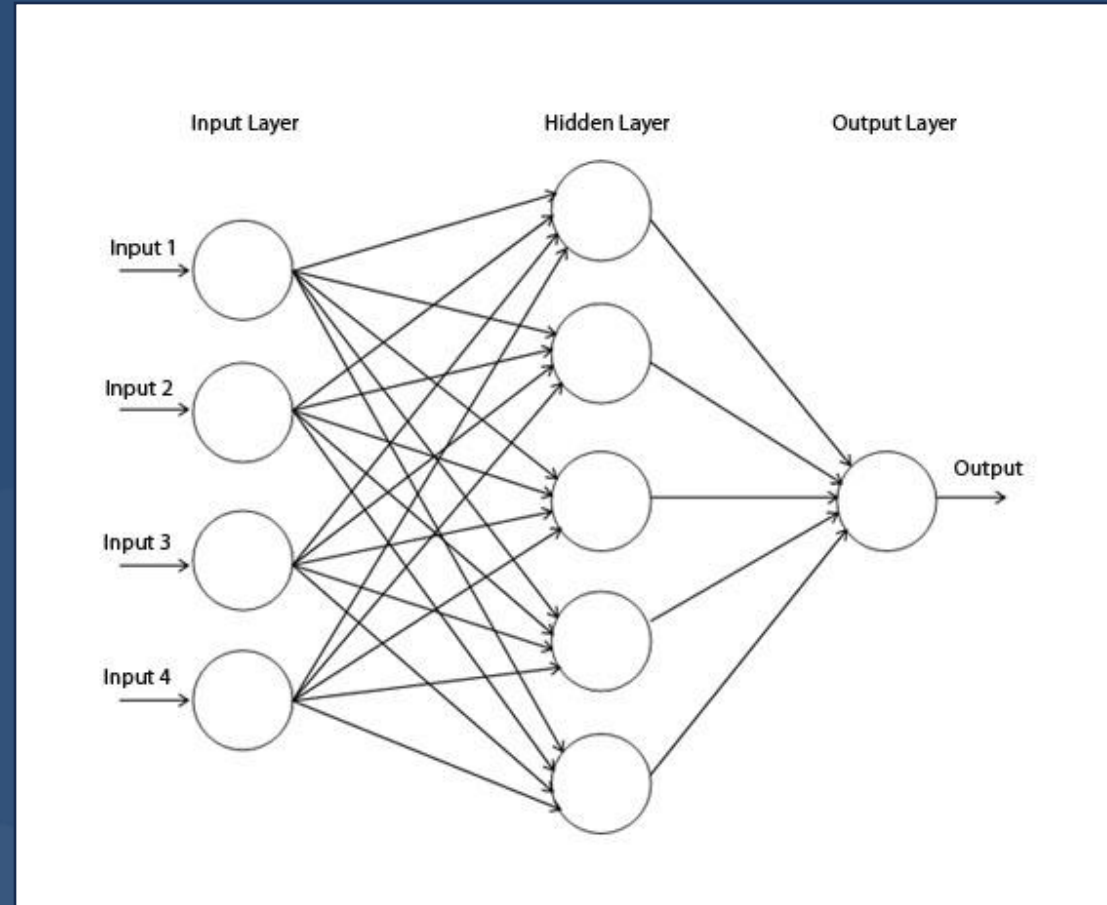
**It's Magic!**  
(thank you for attending ...)



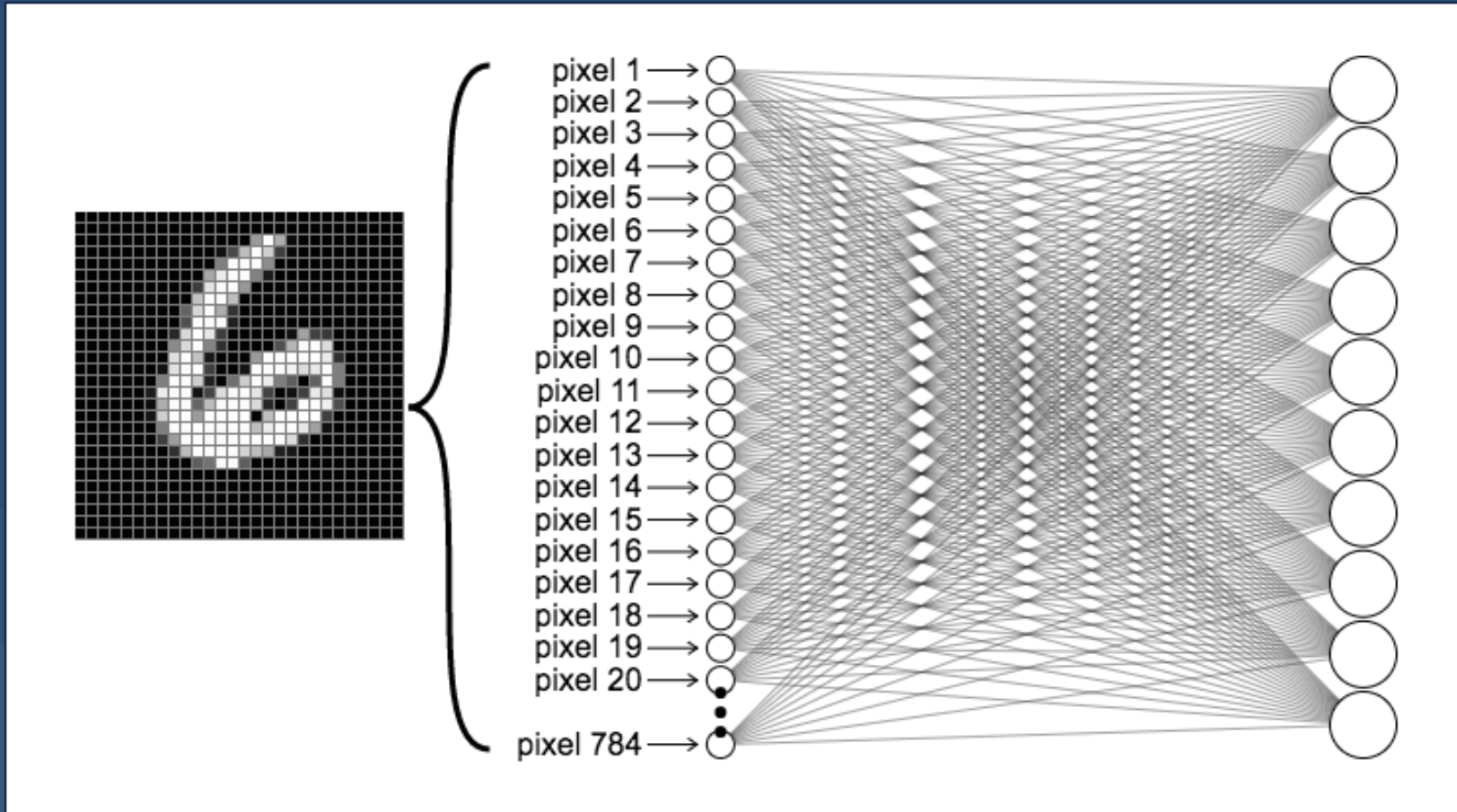
# Segmentation



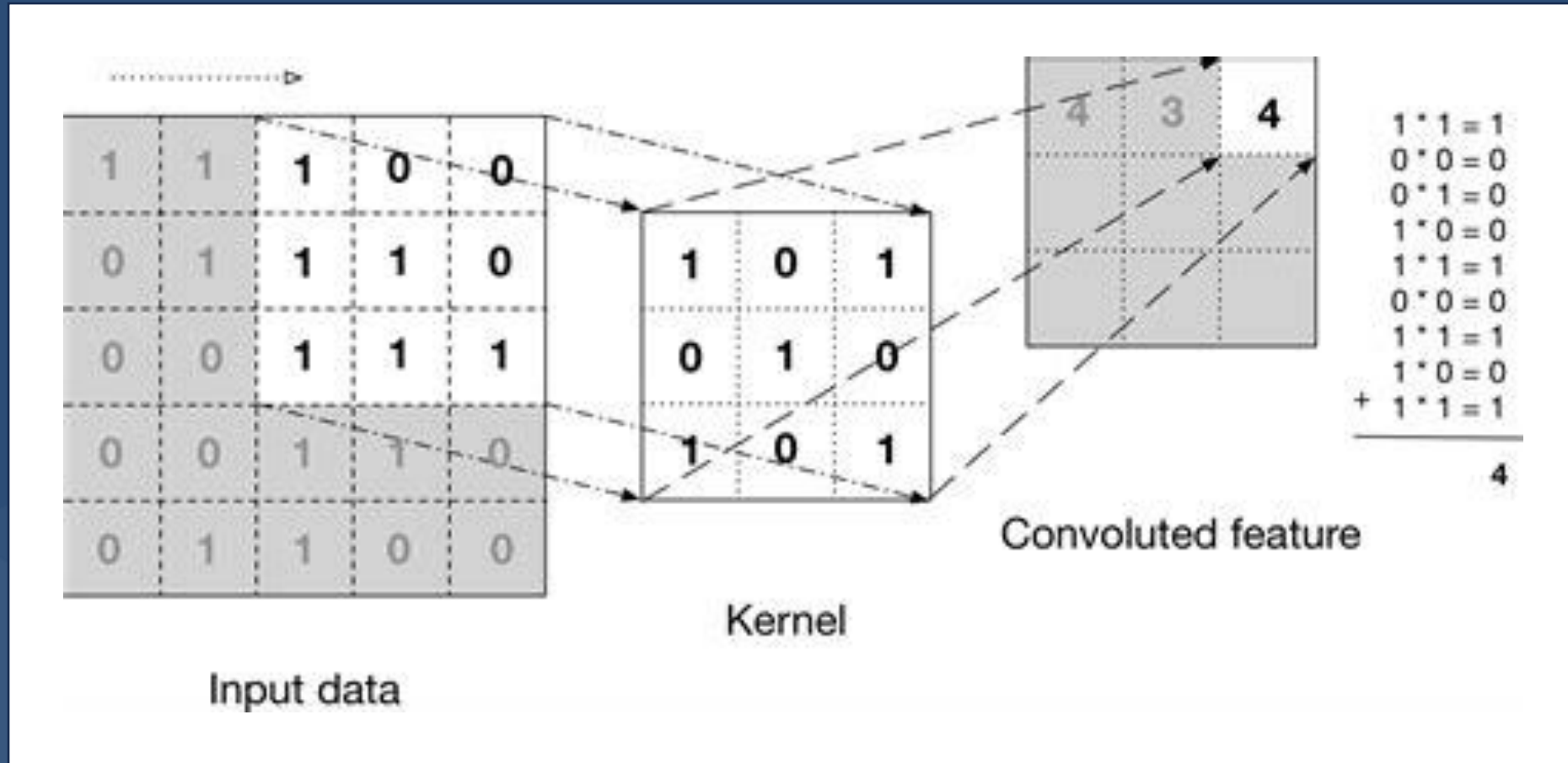
# Fully Connected Layers



# Fully Connected Layers



# Convolutional Layers



from "Deep Learning" by Adam Gibson, Josh Patterson

# Convolutional Layers

CODE  
PRESENTS



Original Image



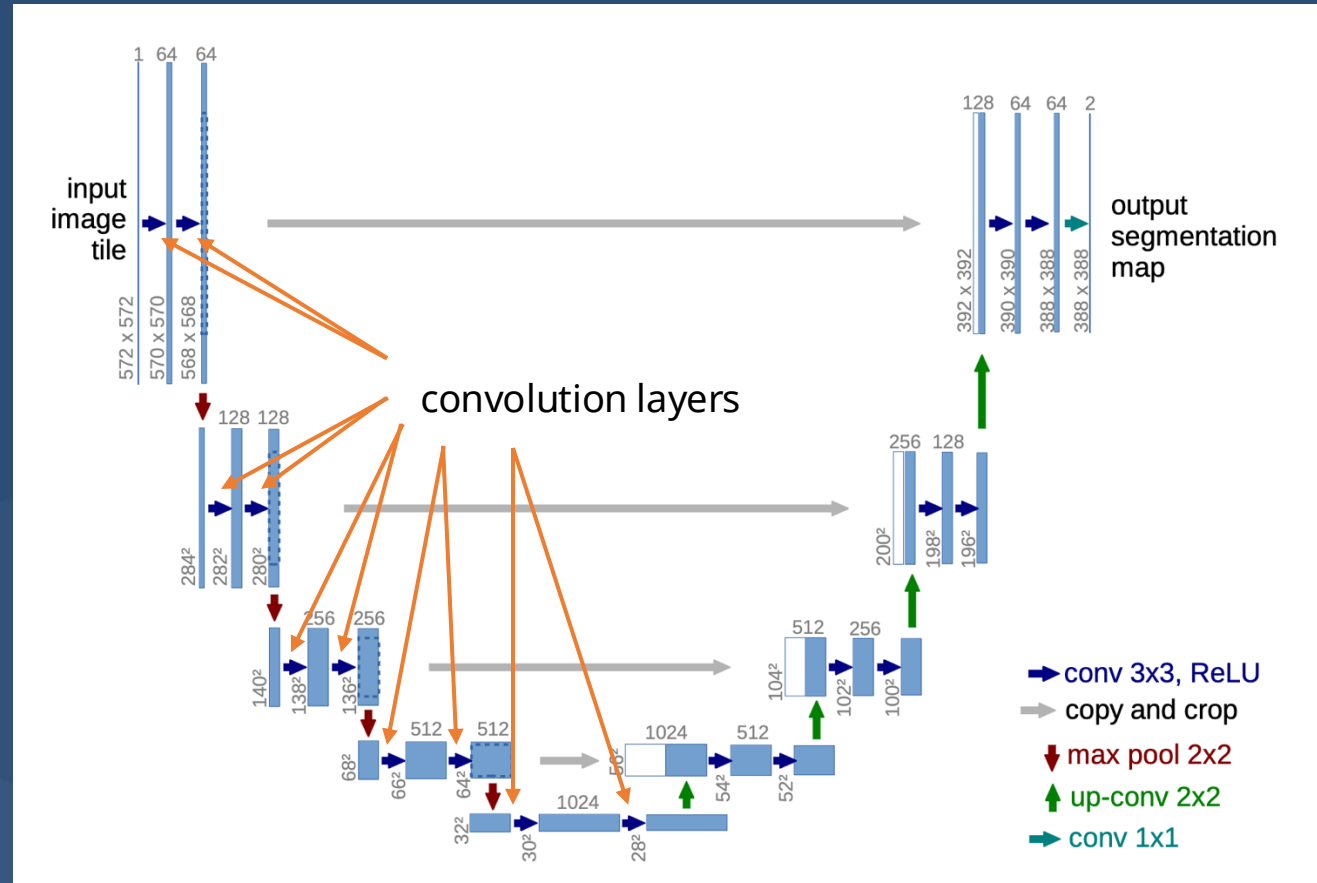
Filters



Segmented Image

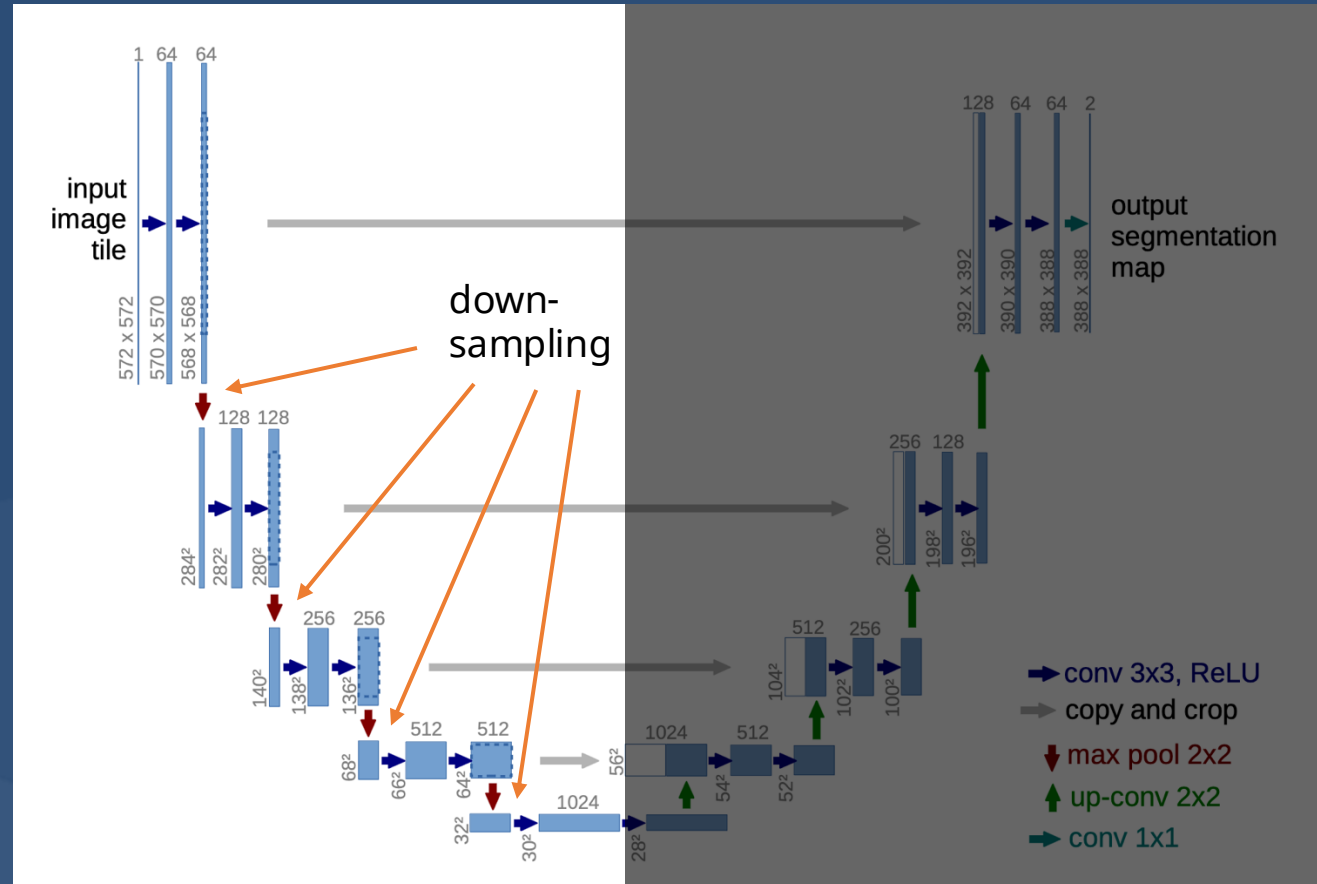


# U-Net Architecture



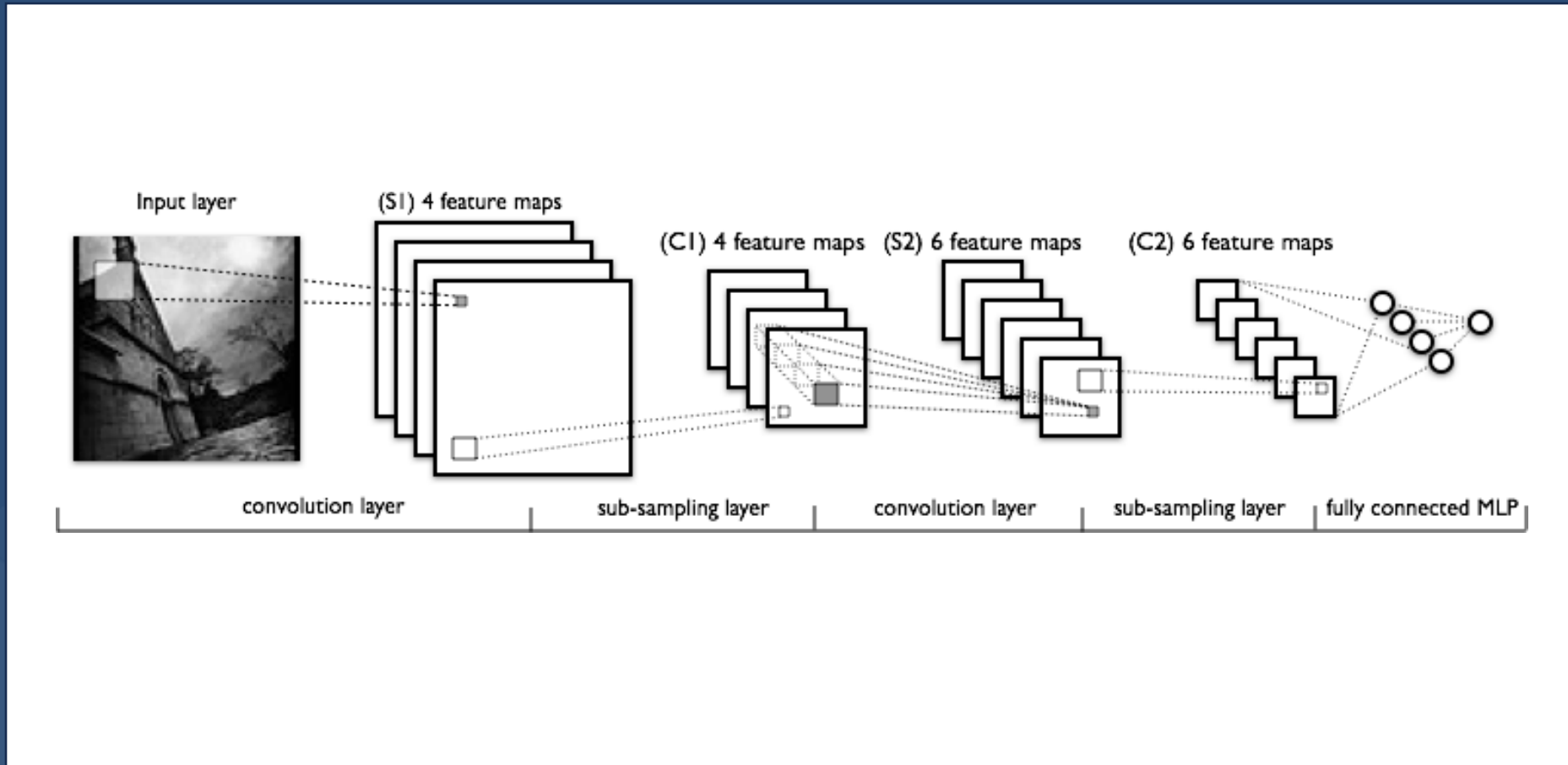
from: **U-Net: Convolutional Networks for Biomedical Image Segmentation**  
arXiv:1505.04597 [cs.CV], Olaf Ronneberger, Philipp Fischer, and Thomas Brox

# Downsampling



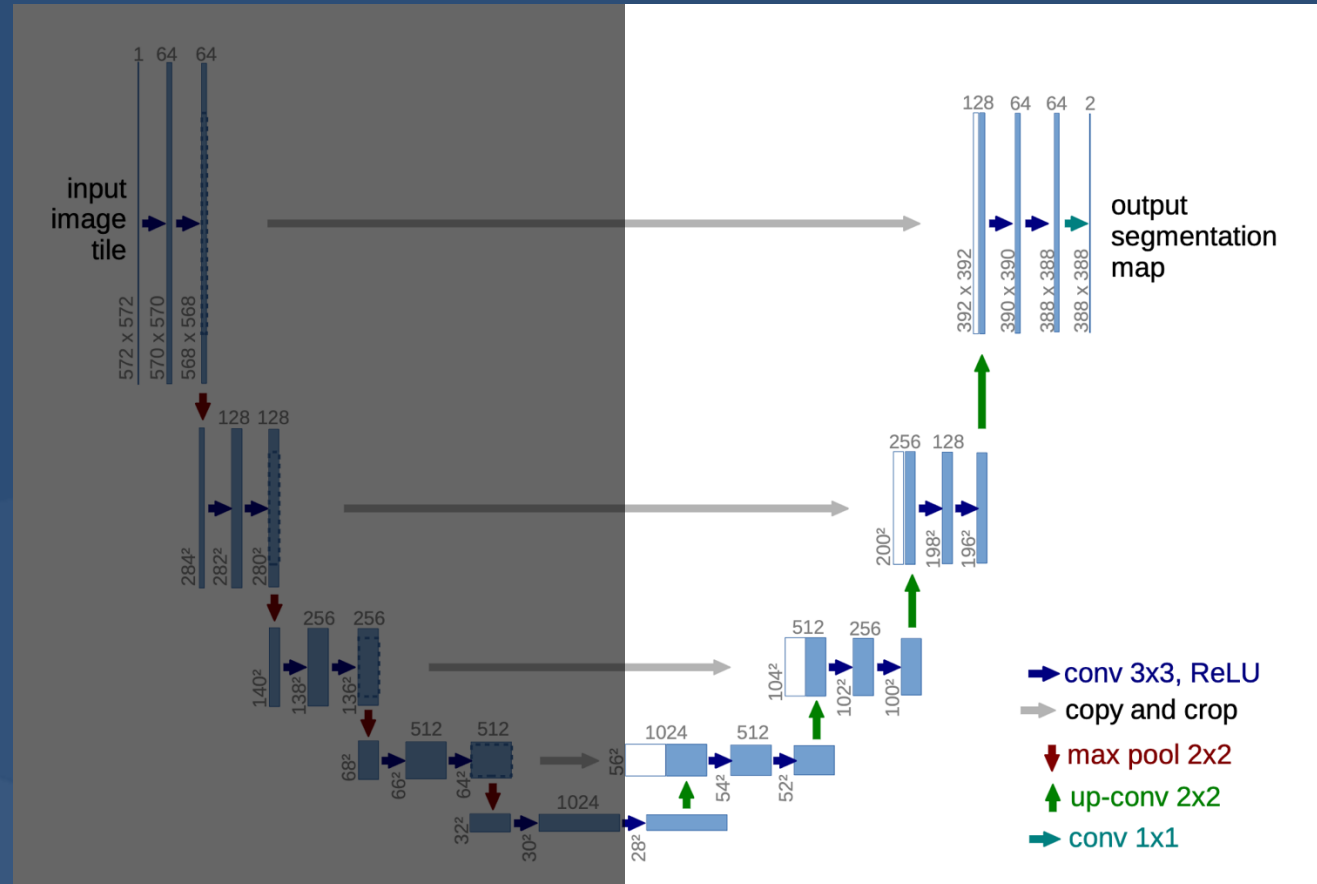
from: **U-Net: Convolutional Networks for Biomedical Image Segmentation**  
arXiv:1505.04597 [cs.CV], Olaf Ronneberger, Philipp Fischer, and Thomas Brox

# Feature Extraction



from <http://deeplearning.net/>

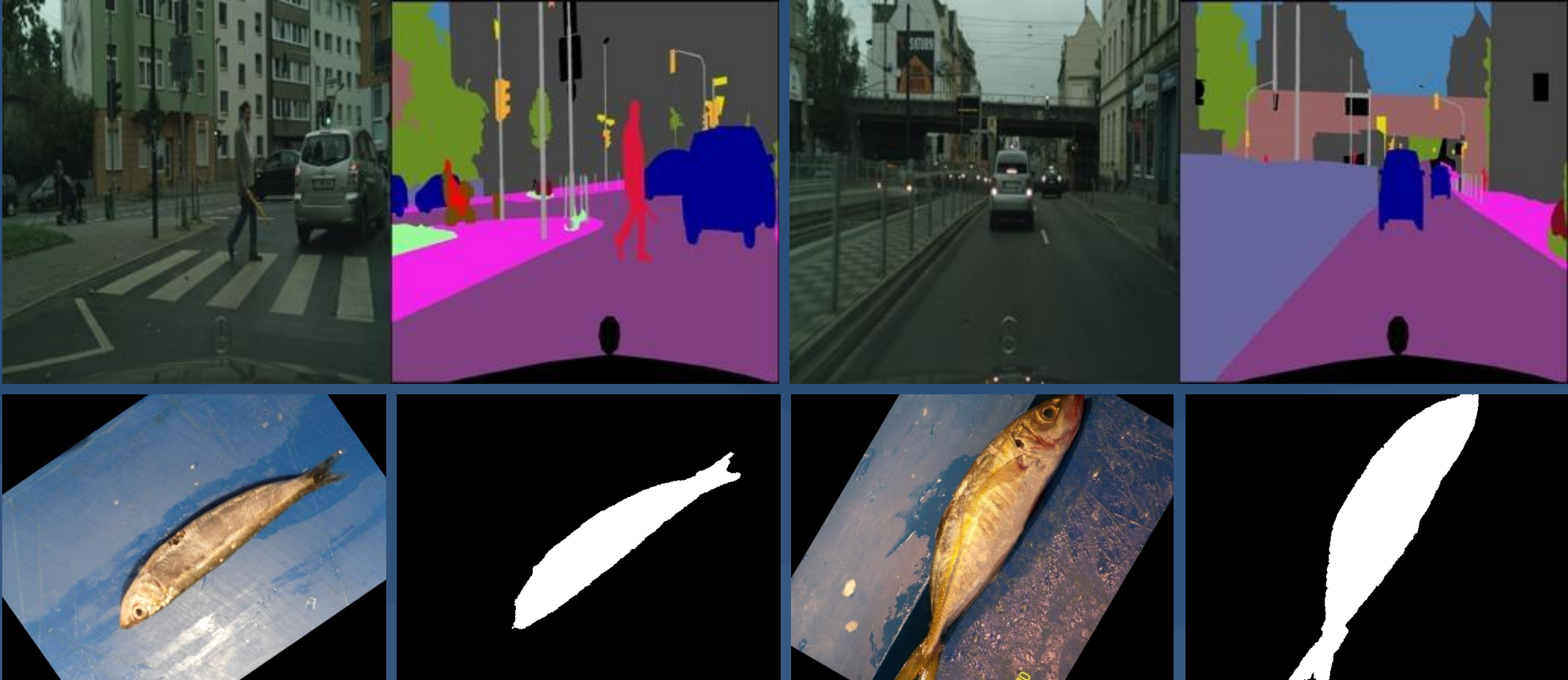
# Upsampling



from: **U-Net: Convolutional Networks for Biomedical Image Segmentation**  
arXiv:1505.04597 [cs.CV], Olaf Ronneberger, Philipp Fischer, and Thomas Brox

# Training

CODE  
PRESENTS



# Diffusion Models



# Denoising with U-Net

**CODE**  
PRESENTS



=



+



# Denoising with U-Net





# Denoising with U-Net



Input



Ground Truth

# Denoising with U-Net



0



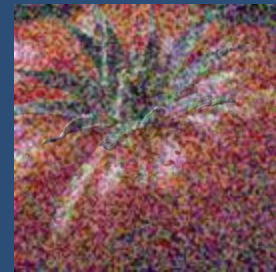
5



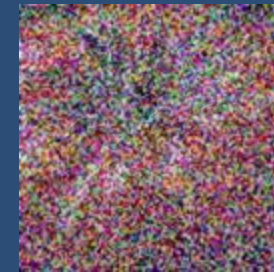
10



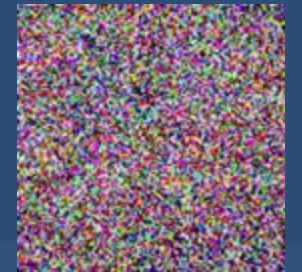
15



20

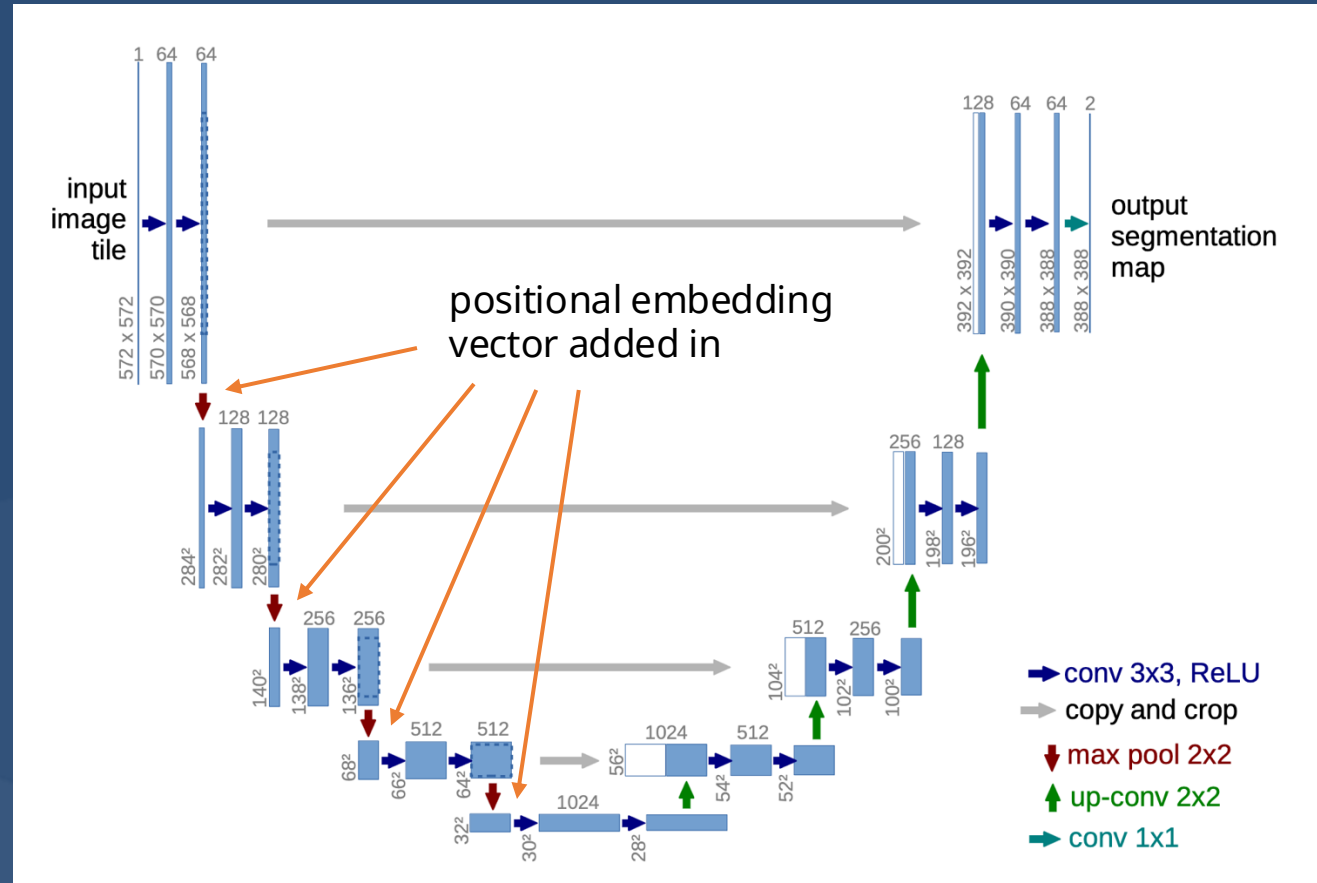


25



30

# Denoising with U-Net

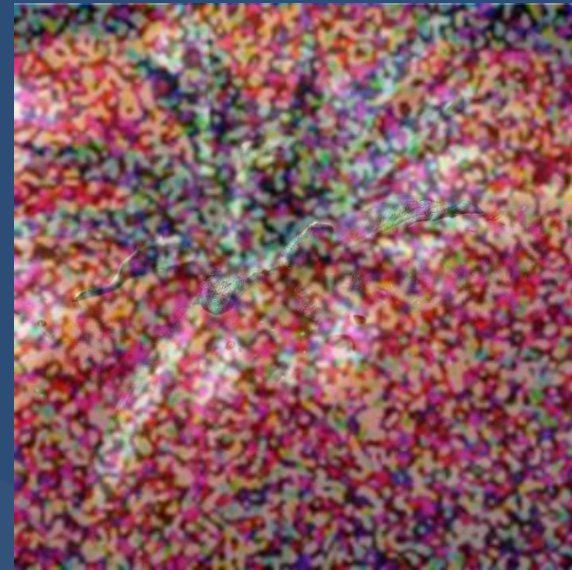


from: **U-Net: Convolutional Networks for Biomedical Image Segmentation**  
arXiv:1505.04597 [cs.CV], Olaf Ronneberger, Philipp Fischer, and Thomas Brox

# Denoising with U-Net



Noise Prediction

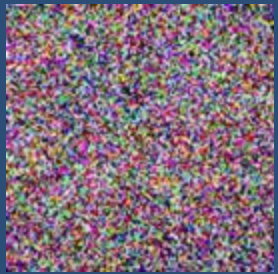


Denoised Prediction



# Denoising with U-Net

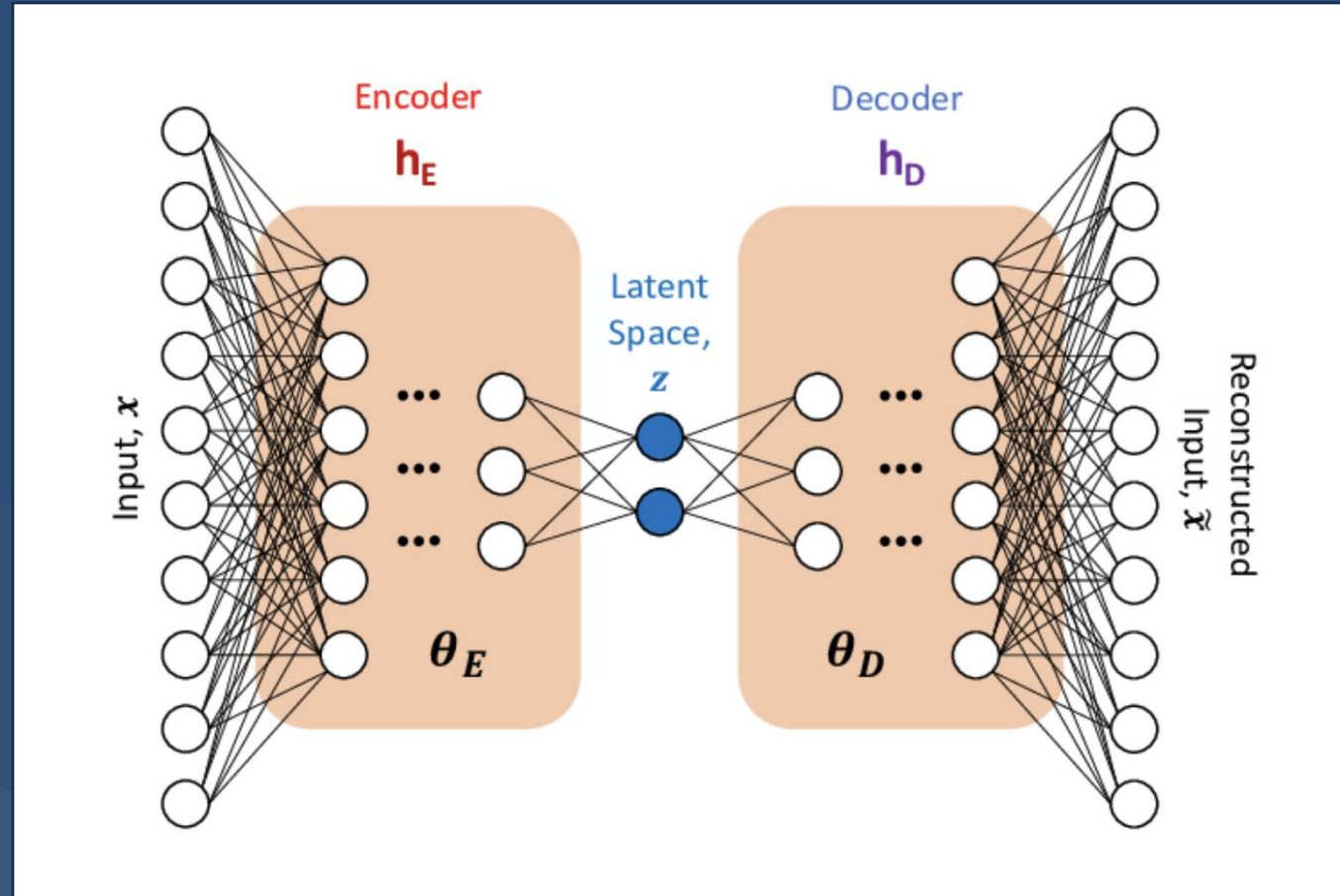
**CODE**  
PRESENTS



# Latent Space / Variational Auto-Encoders



# Variational Auto-Encoders



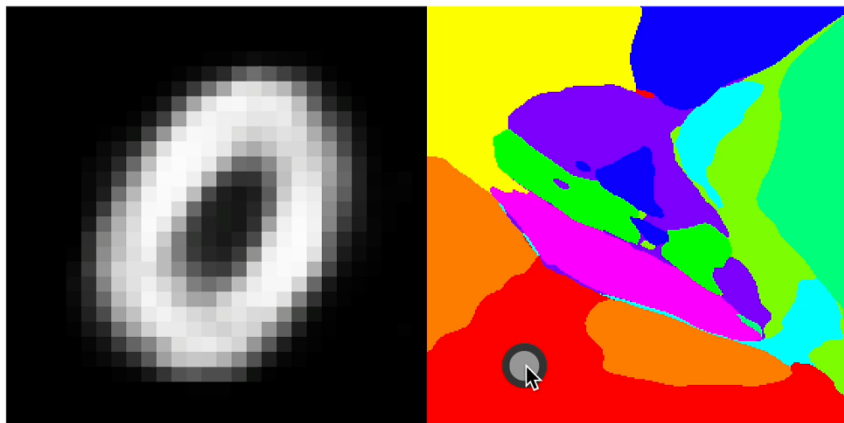
from: Dutta, Sourav & Rivera-Casillas, Peter & Cecil, Orie & Farthing, Matthew & Perracchione, Emma & Putti, Mario. (2021). Data-driven reduced order modeling of environmental hydrodynamics using deep autoencoders and neural ODEs.

# Variational Auto-Encoders

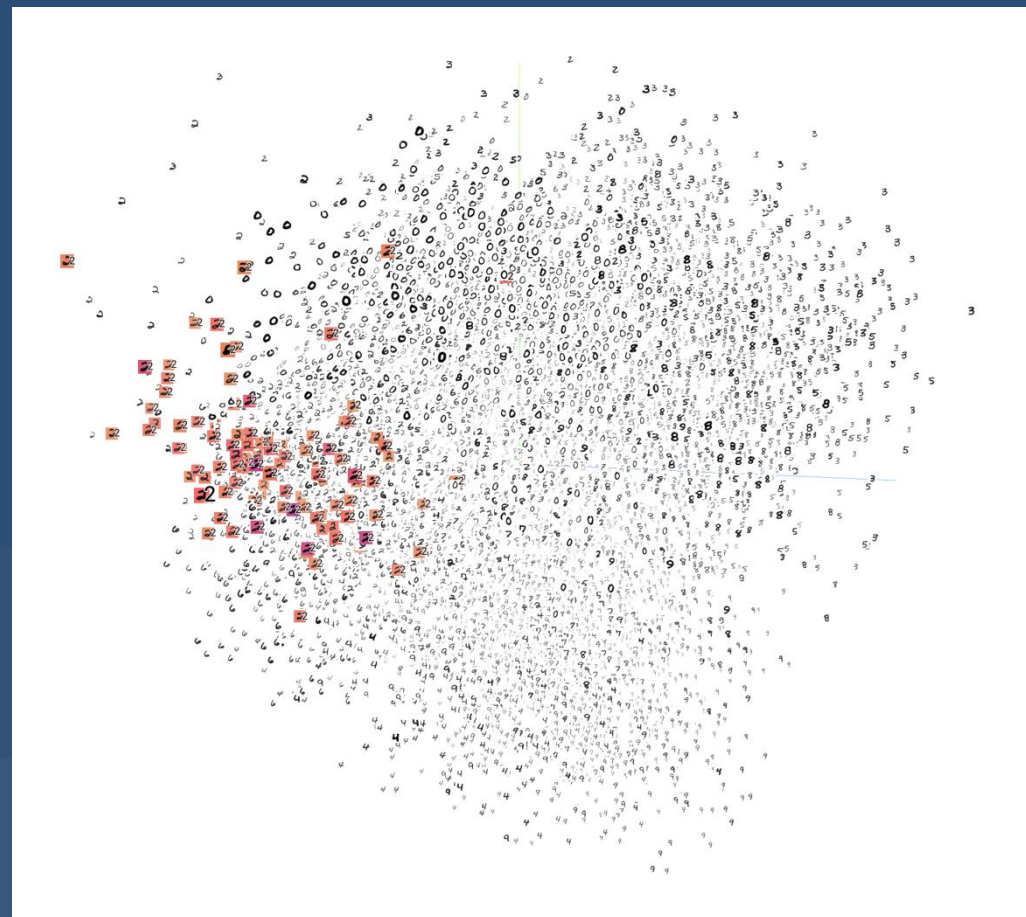
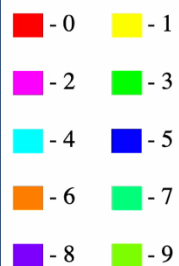
**CODE**  
PRESENTS

Made by using tensorflow.js to train an autoencoder on the MNIST dataset.

Drag the circle around to explore the latent space of digits.



**Legend:**



# Latent Space

**CODE**  
PRESENTS



?





# Latent Diffusion Model

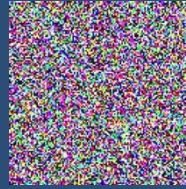
**CODE**  
PRESENTS



▶  
encode

latent  
space

+



-

latent  
space

◀  
decode



# Text to Image Generation

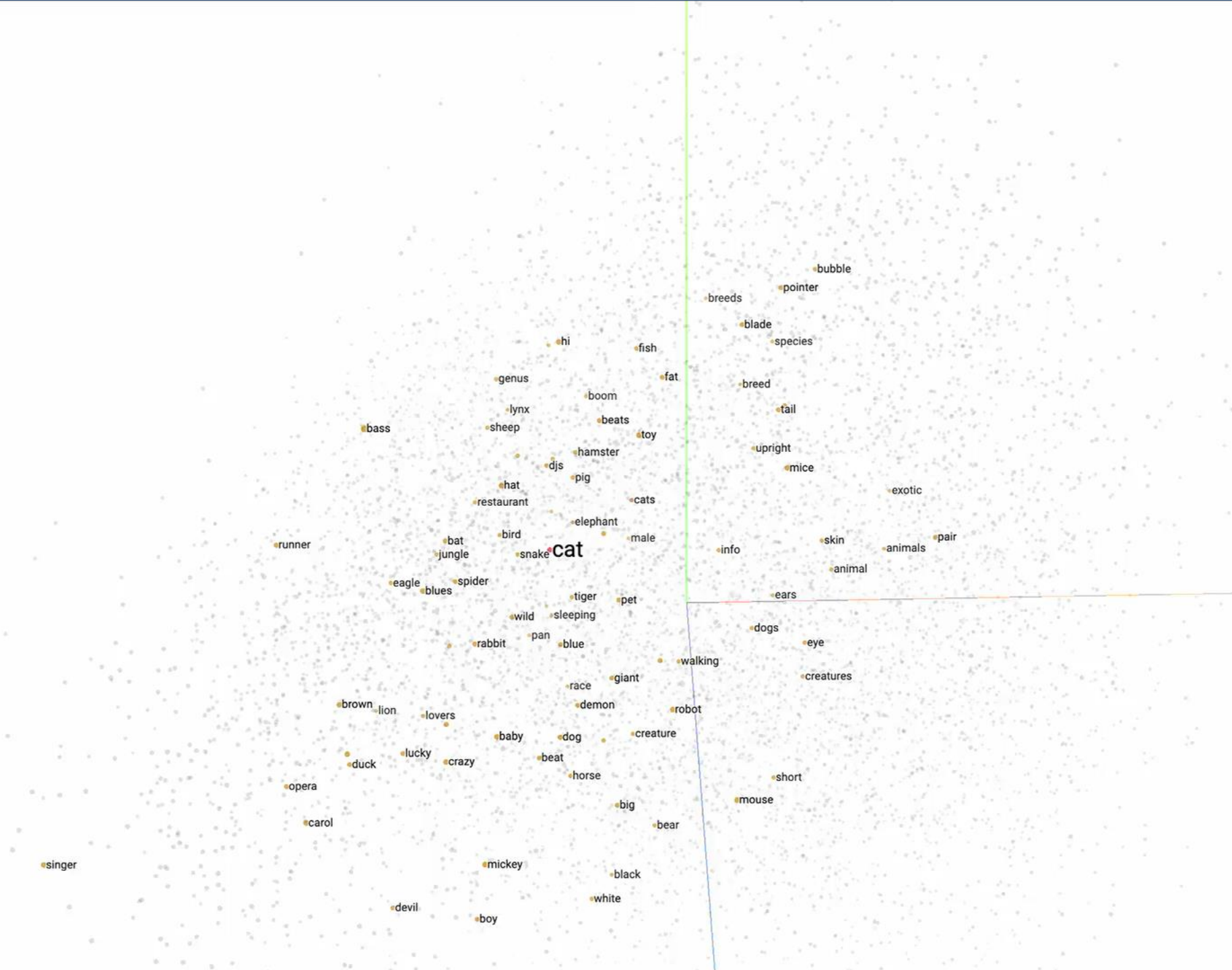


# Text to Image Generation

**CODE**  
PRESENTS

astronaut cat, playing piano in  
space, black hole in the  
background





# Text to Image Generation

CODE  
PRESENTS



convolutional  
encoder

[0.34, 0.74, 0.25, ...]

astronaut cat, playing piano  
in space, black hole in the  
background

self attention  
encoder

[0.94, 0.52, 0.71, ...]

# Text to Image Generation

**CODE**  
PRESENTS



convolutional  
encoder

[0.34, 0.74, 0.25, ...]

astronaut cat, playing piano  
in space, black hole in the  
background

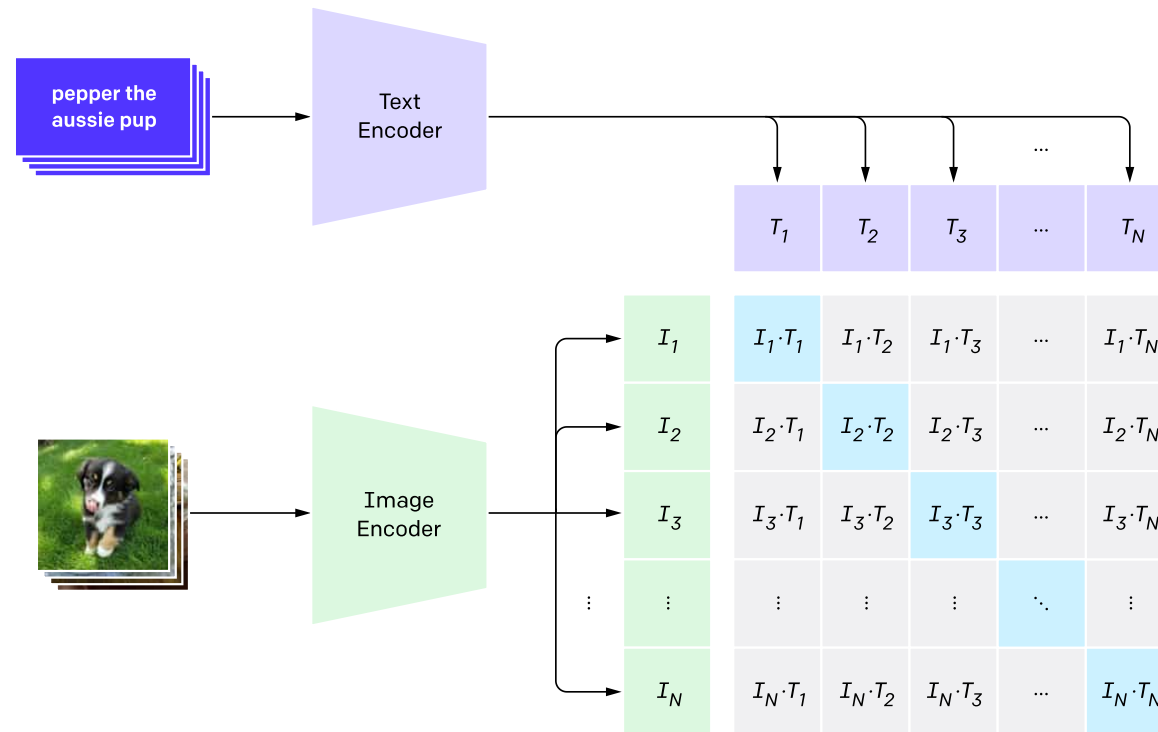
self attention  
encoder

[0.34, 0.74, 0.25, ...]



# Contrastive Language - Image Pre-training (CLIP)

## 1. Contrastive pre-training



from: <https://openai.com/research/clip>

**CODE**  
*PRESENTS*

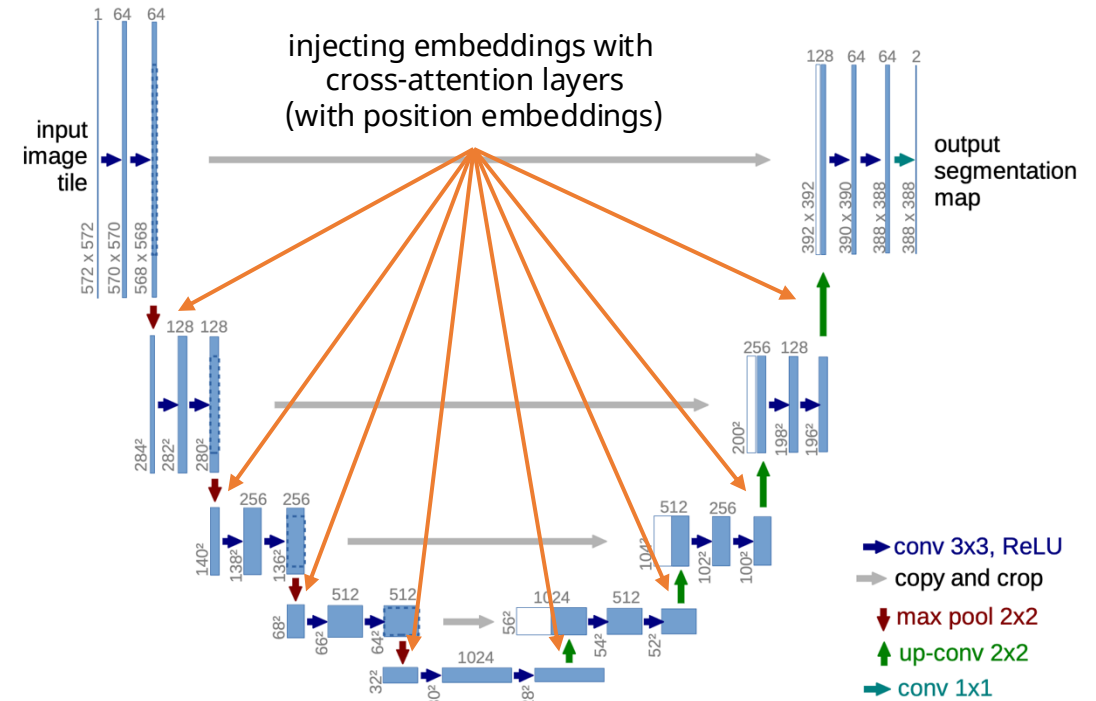
The diagram illustrates a multi-modal model architecture for image captioning. It consists of two main components: a Text Encoder and an Image Encoder.

**Text Encoder:** Takes the input text "pepper the aussie pup" and processes it into a sequence of tokens  $T_1, T_2, T_3, \dots, T_N$ .

**Image Encoder:** Takes a stack of images (showing a dog) and processes them into a sequence of features  $I_1, I_2, I_3, \dots, I_N$ .

**Feature Matrix:** The outputs of both encoders are combined into a matrix where each element represents the dot product of an image feature and a text token. The matrix is structured as follows:

$I_1$	$I_1 \cdot T_1$	$I_1 \cdot T_2$	$I_1 \cdot T_3$	$\dots$	$I_1 \cdot T_N$
$I_2$	$I_2 \cdot T_1$	$I_2 \cdot T_2$	$I_2 \cdot T_3$	$\dots$	$I_2 \cdot T_N$
$I_3$	$I_3 \cdot T_1$	$I_3 \cdot T_2$	$I_3 \cdot T_3$	$\dots$	$I_3 \cdot T_N$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$I_N$	$I_N \cdot T_1$	$I_N \cdot T_2$	$I_N \cdot T_3$	$\dots$	$I_N \cdot T_N$



# Cross Attention

**CODE**  
PRESENTS

astronaut cat, playing piano in  
space, black hole in the  
background





# Cross Attention

CODE  
PRESENTS

astronaut cat, playing piano in  
space, black hole in the  
background



A photograph showing a group of police officers in uniform surrounding a man in a suit, likely Donald Trump, in a crowd. The scene is dimly lit, and the image has a dark, moody tone. The text "Impact of Generative AI" is overlaid in the center.

# Impact of Generative AI

# Considerations

- Collection of Training Material
- Biases in Training Data (racial, gender, cultural)
- Responsible Product Integration
- Copyright Issues
- Absence of Legislation / Settled Case-Law
- ...



# Loss of Reputation

**CODE**  
PRESENTS

## LEGO caught using artificial intelligence (AI) to generate fake Ninjago images

March 15, 2024 | 18



SAMSUNG / TECH / CAMERAS AND PHOTOGRAPHY

## Samsung caught faking zoom photos of the Moon



/ A viral Reddit post has revealed just how much processing the company's cameras apply to photos of the Moon, further blurring the line between real and fake imagery in the age of AI.

RETAIL

## Instacart is using AI art. It's incredibly unappetizing.

Jake Swearingen

Share Save



Photos of food on Instacart that appear to be generated by AI. Instacart/Business Insider

## Wizards of the Coast admits using AI art after banning AI art

Magic: The Gathering publisher initially insisted the marketing image was human-made

BY OLI WELSH  
Jan 8, 2024, 8:40 AM CST

67 Comments (67 New)



Image: Wizards of the Coast via X

## Sports Illustrated Published Articles by Fake, AI-Generated Writers

We asked them about it – and they deleted everything.

/ Artificial Intelligence / AI / Artificial Intelligence / Journalism



Image by Joe Randle via Getty / Futurism

# Conclusion

- Capabilities and limitations
  - Trainability
  - Text handling
  - Prompt adherence
- How diffusion models work
  - Instance segmentation using U-Net
  - Denoising with U-Net
  - Latent space and VAE
  - Cross Attention
- Impact of generative AI
  - Deepfakes
  - Disinformation
  - Loss of reputation



Subscribe and “ring that bell”  
to never miss any of our content!

[youtube.com/codemag](https://youtube.com/codemag)



YouTube



Home



Shorts



Subscriptions



Library



CODE Magazine

@Codemag 2.91K subscribers 69 videos

For over 20 years, CODE Magazine has provided technical content in our pr... >

Subscribed ▾



# Free Subscription

**CODE**  
MAGAZINE

The leading software development magazine,  
written by expert developers for developers.

**Free subscription for all of you**  
(and your friends)

Subscribers get CODE Focus  
issues free of charge!

Share this link to our free subscription:  
**[bit.ly/CP121824Subscribe](https://bit.ly/CP121824Subscribe)**



# Event Survey – Win \$100!



Complete this short 12 question survey for a chance at a **\$100 Amazon Gift Card!**

Survey must be completed by  
**11:59pm ET on Friday 11/01/2024**  
to be eligible!

THIS SLIDE WILL BE REPEATED AT THE END  
AND SURVEY LINK REPEATED IN THE CHAT WINDOW!

**[bit.ly/CP121824Survey](https://bit.ly/CP121824Survey)**



## CODE Presents: Prompt Engineering Talk to an AI Survey

The survey will take approximately 4 minutes to complete.

Thank you for attending! Please complete this brief 12 question survey.  
Your survey must be completed by 11:59pm ET (UTC-4) on Friday, November 1st, 2024.  
Please. Drawing will occur and the individual winner notified.

Thank you for attending! Please complete this brief survey.  
If you cannot attend but watched the recording instead.

\* Required

1. Full Name \*

Enter your answer

2. Company Name \*

Enter your answer

# Q&A

## Contact us with questions!

### CODE Contact

[codemag.com](http://codemag.com)  
[info@codemag.com](mailto:info@codemag.com)  
[facebook.com/codemag](https://facebook.com/codemag)  
[twitter.com/codemagazine](https://twitter.com/codemagazine)

### Presenter Contact

[pbauer@codemag.com](mailto:pbauer@codemag.com)  
[linkedin.com/in/philippjbauer](https://linkedin.com/in/philippjbauer)  
[philippbauer.bsky.social](https://philippbauer.bsky.social)