

Contextual Precision. Unleashing Super Resolution with Transformers



Oguzhan Cagirir

Department of Informatics
King's College London

About Me

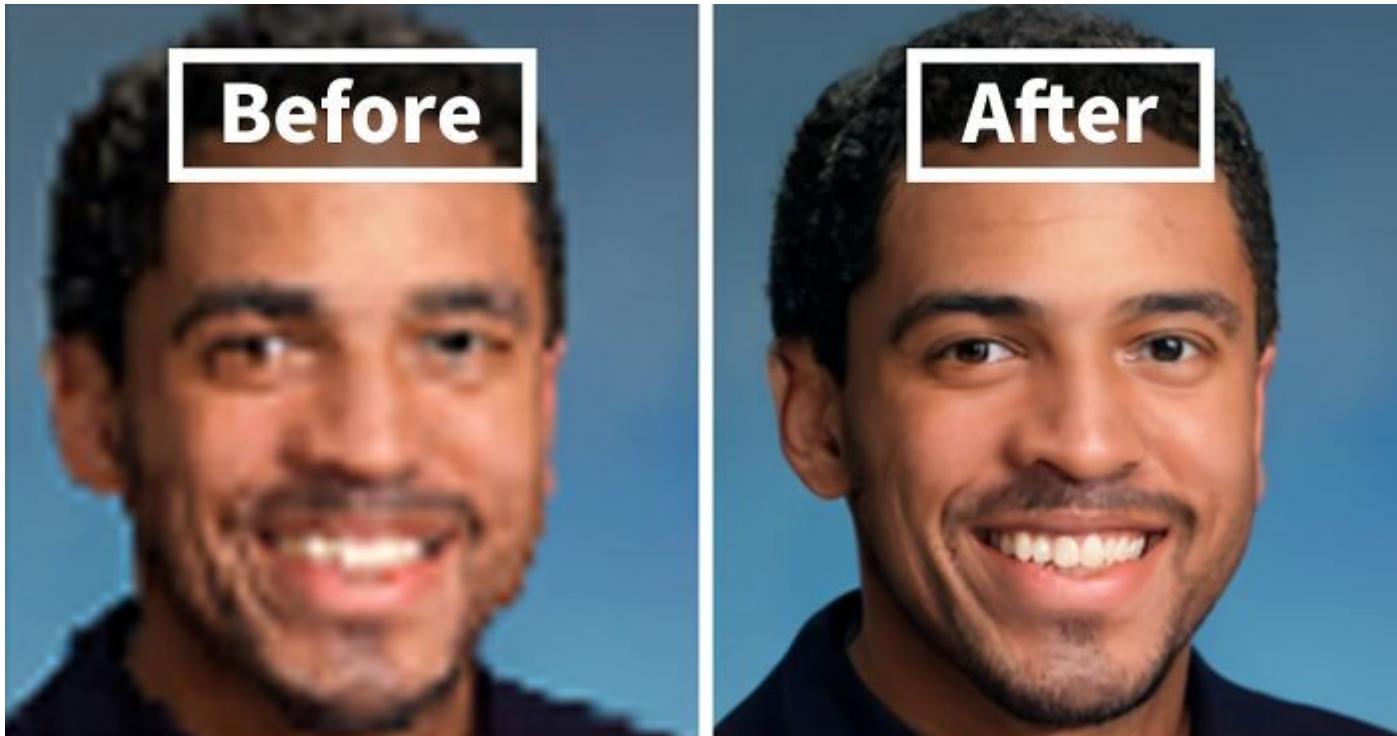
- I'm Oguzhan. People call me 'Oze', so feel free to do so!
- First year undergraduate; Artificial Intelligence @ KCL.
- Full-time Londoner.
- Research into quant & computer vision / SISR.
- Total hardware/gearhead!



PREFACE

So what is Super Resolution?

Super Resolution! (SR)



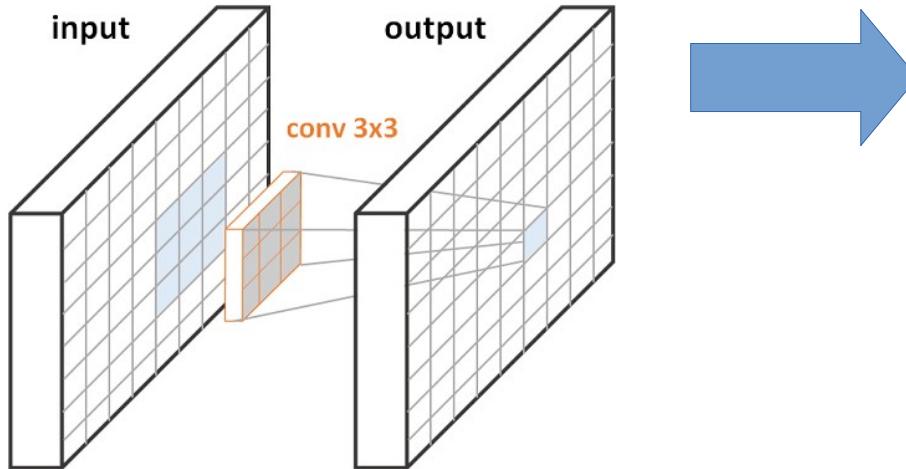
Bad quality → High quality

However, there is no definitive method!

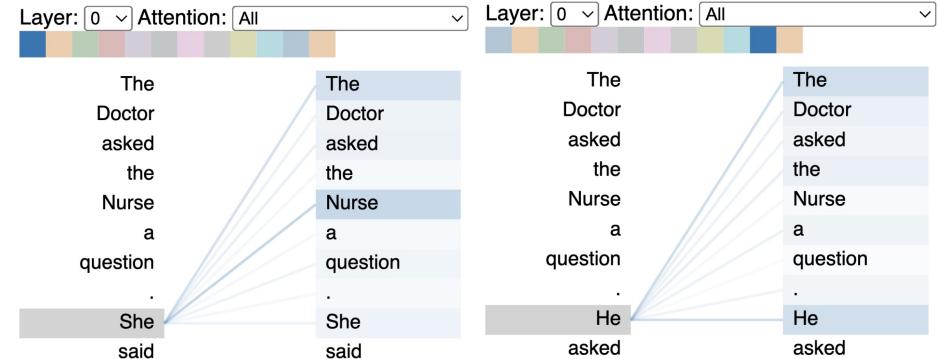
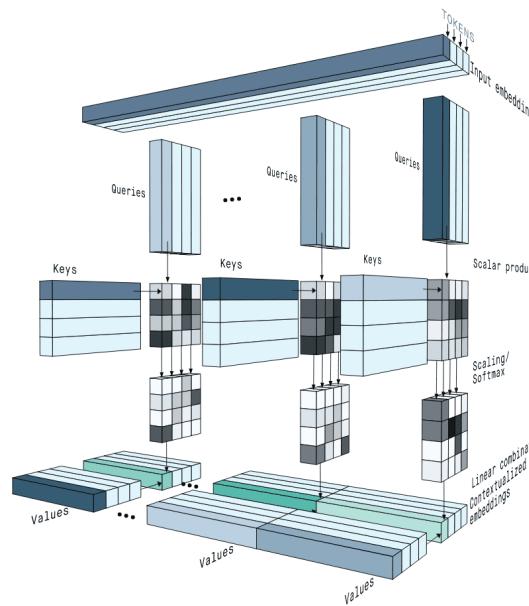
The SR Revolution

Convolutional Networks / CNNs (The old way)

Localised context.

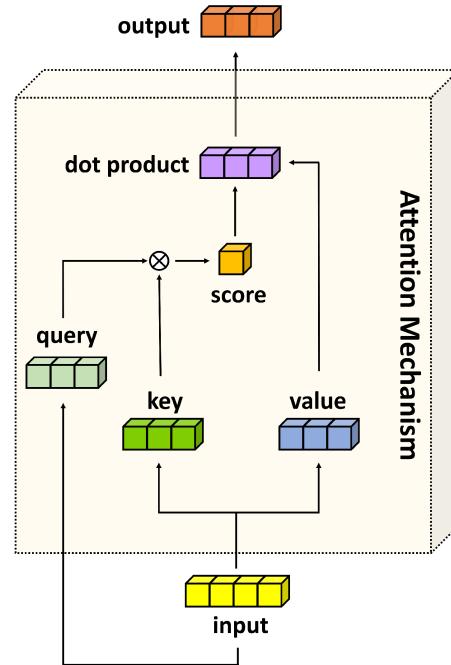
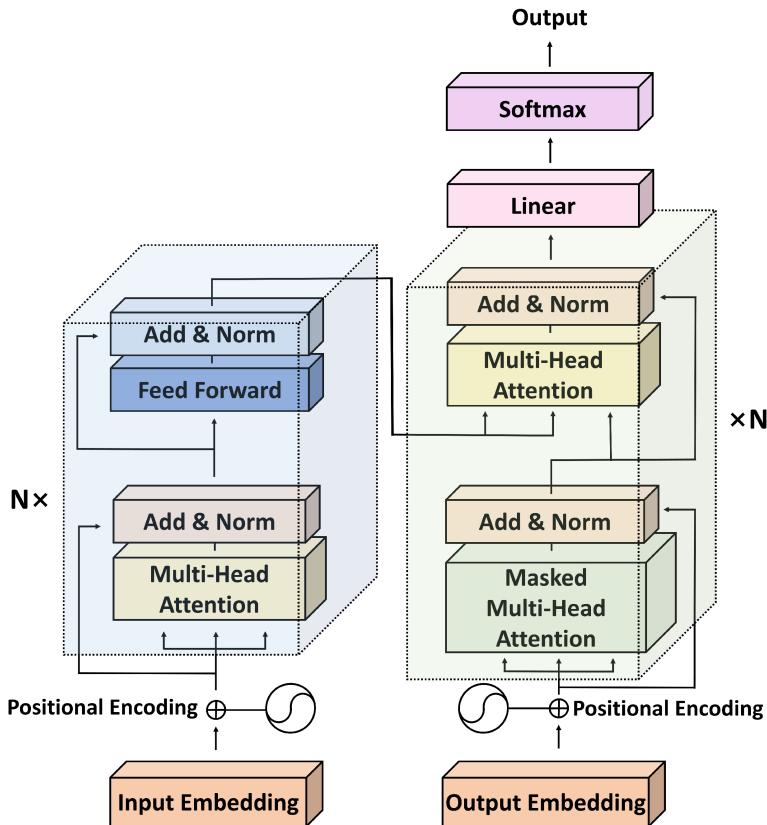


Transformer-based models



It's all about mapping global context & attention!

Self-Attention. Where it's all at...



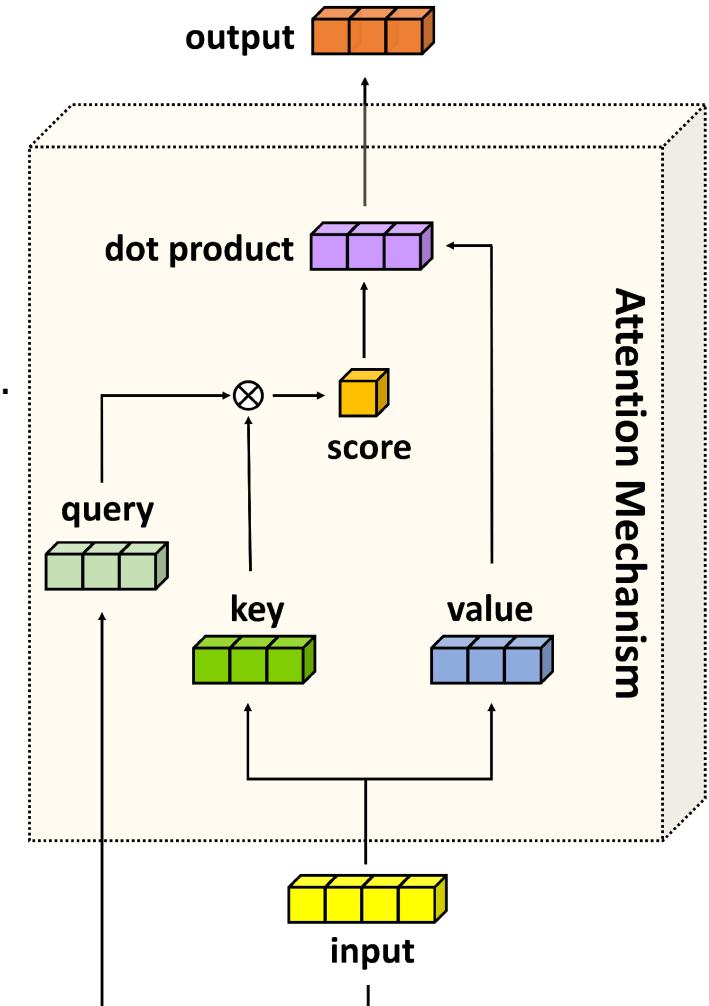
Self-Attention. Where it's all at...

Transformers are primarily suited for text... (LLMs, etc.)

The 'T' in GPT is 'Transformer' after all!

Where the key concept is: Context-Aware Representation.

But with images... What do we do!?



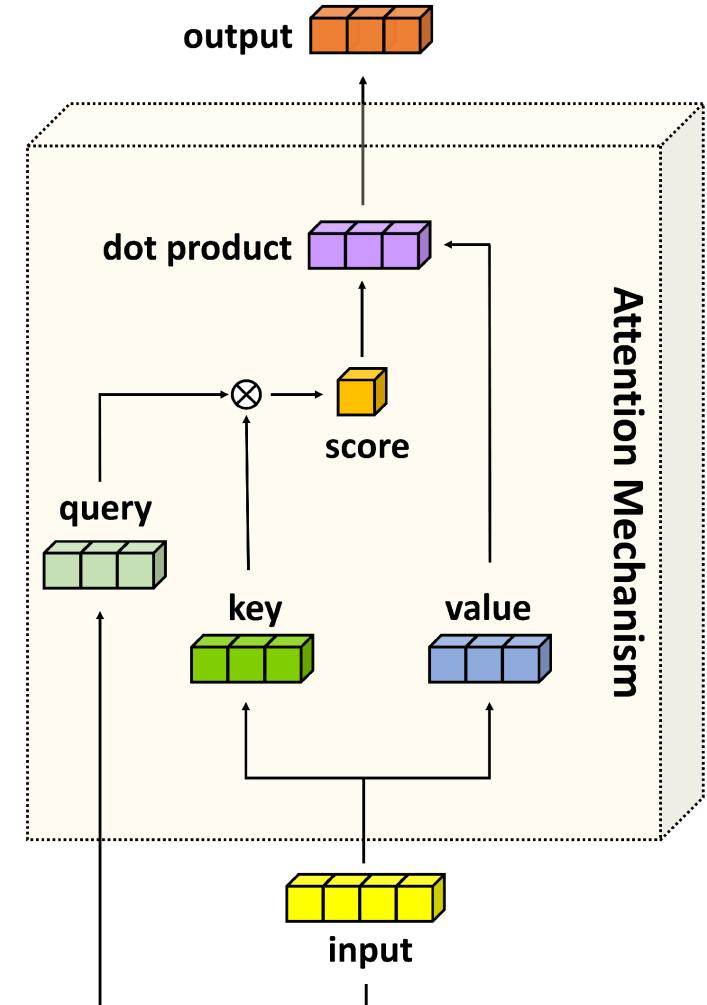
Self-Attention. Where it's all at...

Solution: Break down the image into chunks, and treat them like words.

We know attention works for textual tokens, so this should also work in the same way!

This is the exact premise of a 'Vision' Transformer.

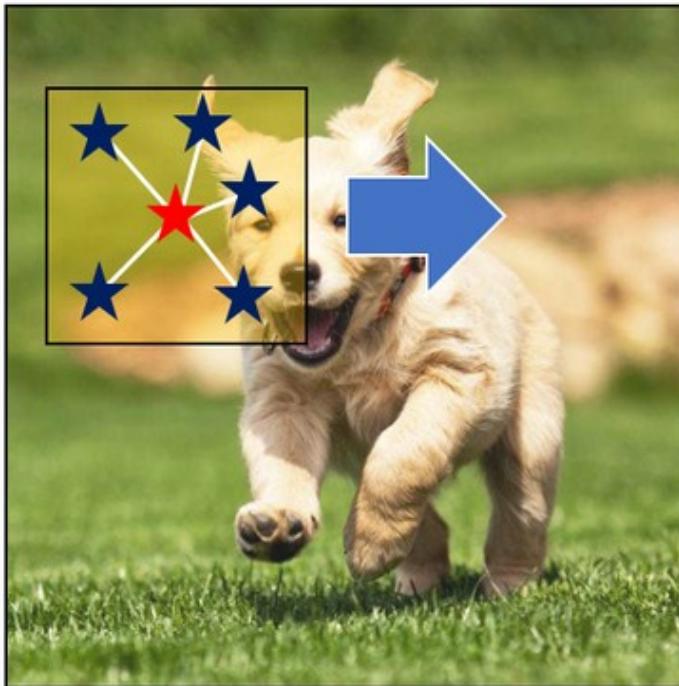
(For those of you familiar with LLMs, those 'chunks' = tokens!)



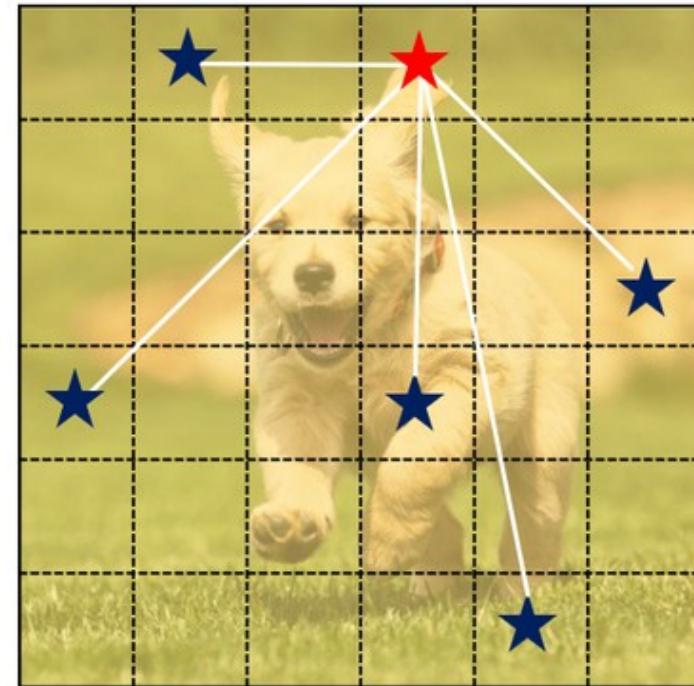
'Chunks' - A lot of them!



Receptive Field



Convolution of CNN

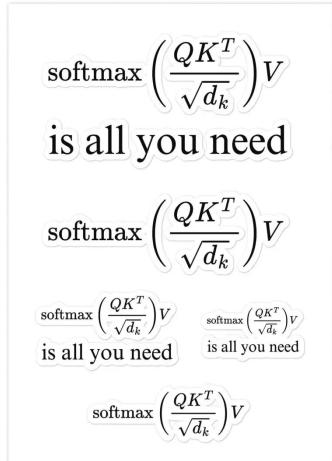


Attention of Vision Transformer

Self-Attention. Again...

But how does any of this actually work anyway?

Far too complicated for starters!



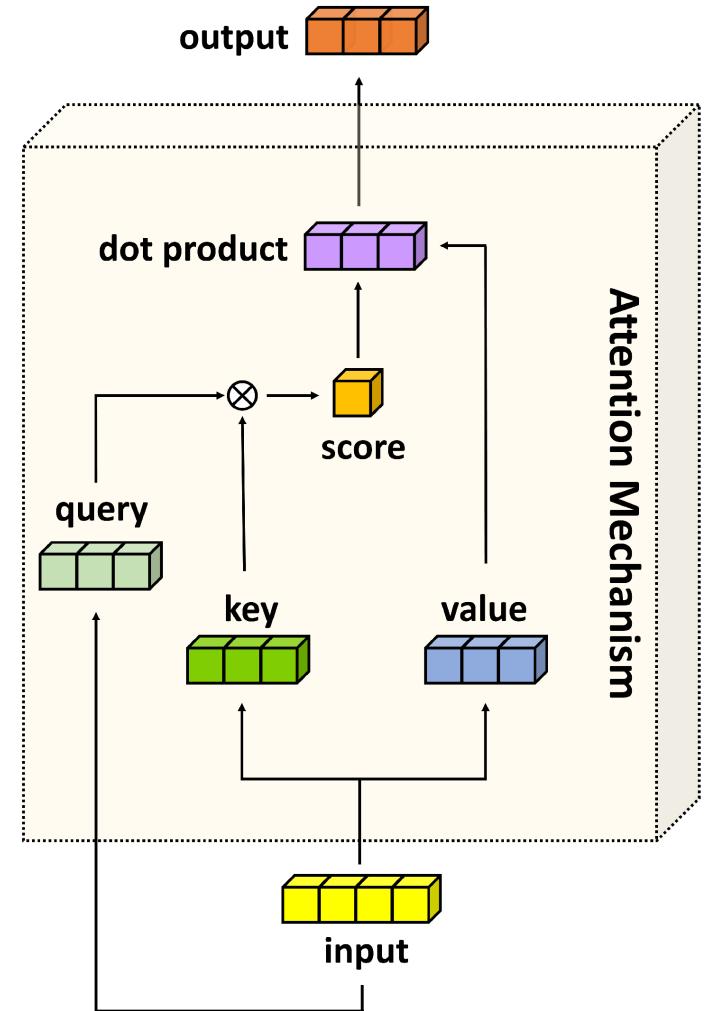
Not helpful

Money Is All You Need

Nick Debu
Tokyo Institute of Bamboo Steamer

Abstract

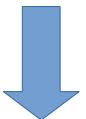
Transformer-based models routinely achieve state-of-the-art results on a number of tasks but training these models can be prohibitively costly, especially on long sequences. We introduce one technique to improve the performance of Transformers. We replace NVIDIA P100s by TPUs, changing its memory from hoge GB to piyo GB. The resulting model performs on par with Transformer-based models while being much more "TSUYO TSUYO".



Self-Attention. Again...

We need to make the model understand the meaning of each 'chunk' in the image.

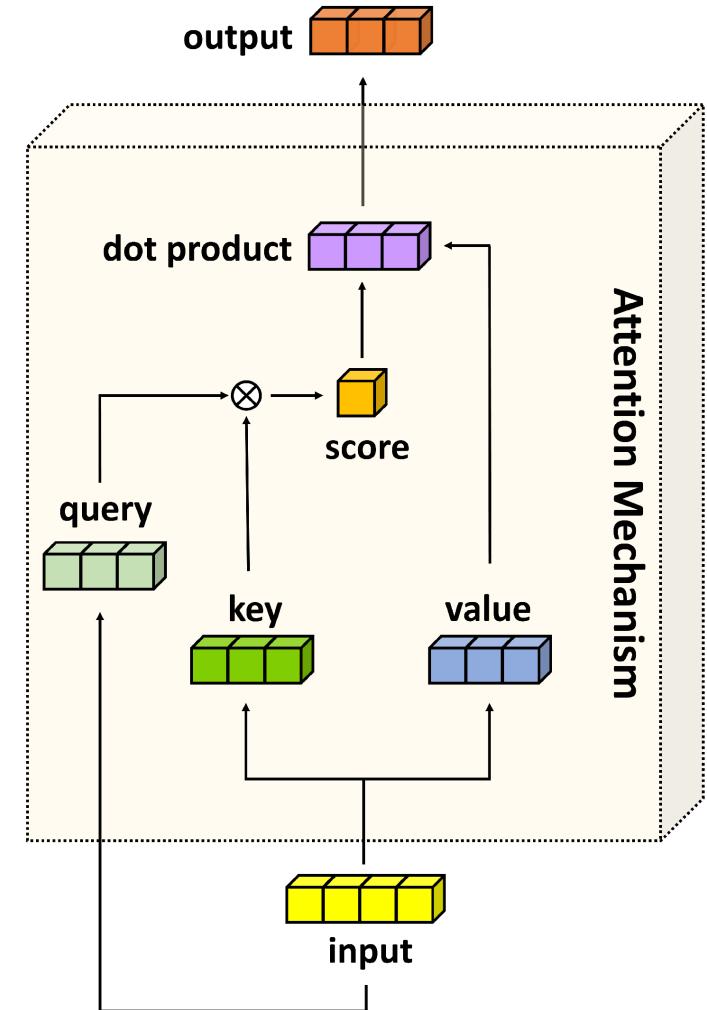
- Query: 'What am I looking for in this image?'
- Key: 'What features do different parts of the image offer?'
- Value: 'What information does parts of the image contain?'



(Query * Key)

- Score: how relevant is each part of the image to the current part we're focusing on?

The result? A model that truly 'sees' the image holistically, understanding context in a way CNNs simply can't.



Technical bit over.

"I understand Oze, but why should I care?"

DLSS 4

Supreme Speed. Superior Visuals.
Powered by AI.



NVIDIA DLSS

Everyone who plays games knows it!

If you haven't, you're missing out.

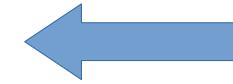




48MP
Fusion camera

24 mm focal length
2.44 µm quad-pixel
100% Focus Pixels
 $f/1.78$ aperture
2x Telephoto at 48 mm
Anti-reflective lens coating
2nd-generation sensor-shift OIS

Explicit SR features
in Galaxy land, with
“Zoom” feature.



Apple doesn't do SR, but
they use similar networks
for post-processing.

One more thing

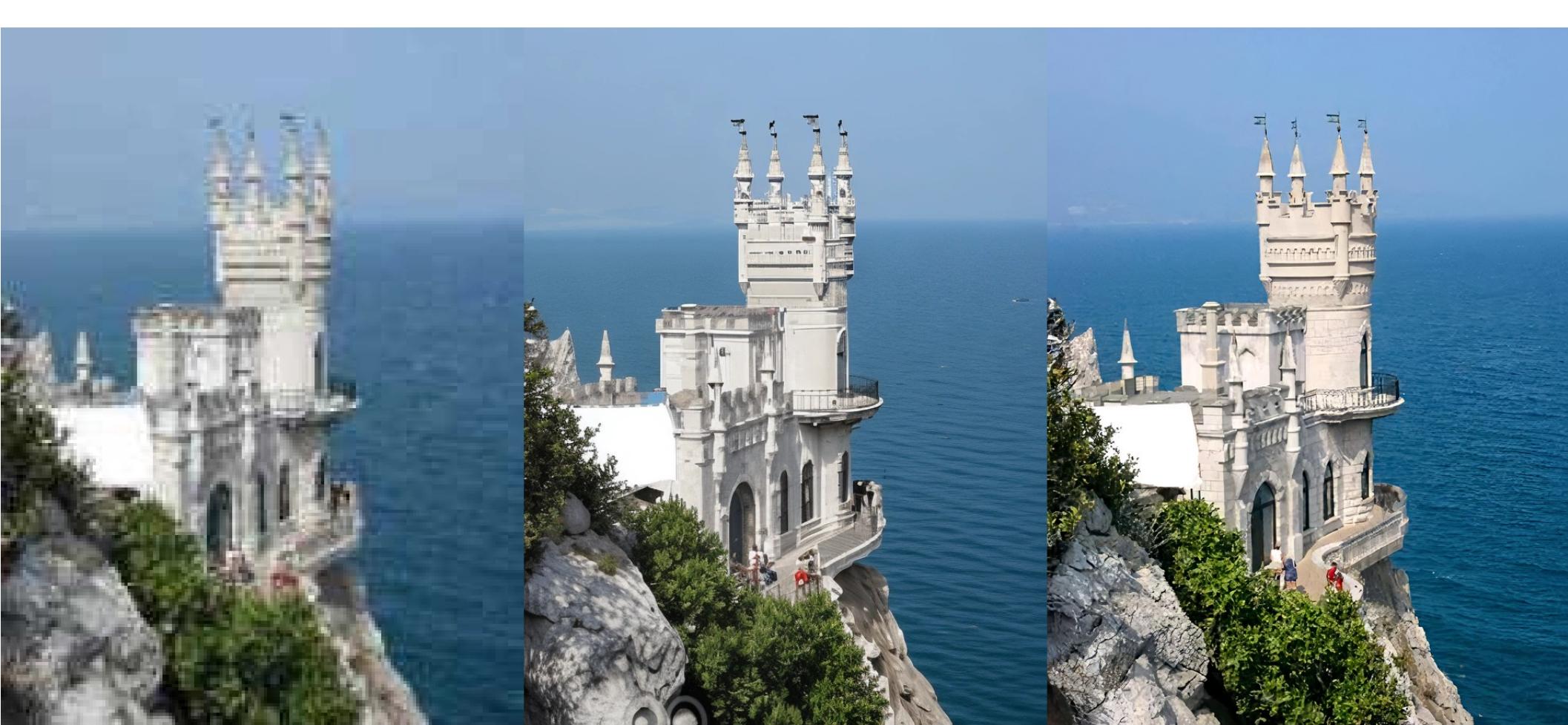
EnhanceFX™ Series – The Hyper-Quality, Super Resolution Architecture.

- Custom Attention Head:
“extremeHEAD”. (What we learned today)
- Ultra-sharp, practically *noiseless*
Hybrid Transformer core.
- Dynamic activation function:
“ViRAU”. (A secret in the making for 5+ years!)
- Quantum-based colour processor. (YES! *REAL* qubits)
- *Full* HDR conversion. (/w PANTONE colour pairing)
- *Video* frame rate interpolation. (24 fps → 384+ fps)
& much more!



Full reveal trailer coming soon!





zsyOAOA / InvSR

EnhanceFX™ VI
QuantumPowerHDR



Gigapixel™ 8

EnhanceFX™ VI
QuantumPowerHDR

Live Demo

Surprise!

Time to put this to the test.

First one to get their phone out and type out this link gets to upload a bad quality image that I'll enhance, **live**.

First come. First serve!

-- LINK REDACTED --

T's&C's: Only upload really low-resolution images - or it won't be impressive!

Links, Papers & Resources (& models)



Everything you need, one place

Wrap-up & Socials

Have an awesome project? Let's talk!

Pinned

EnhanceFX Public EnhanceFX™ is a novel, state-of-the-art hybrid super-resolution model architecture. ⭐ 2

NeuralFusionII Public NeuralFusionII is a custom transformer LLM model crafted using PyTorch, designed for enthusiasts.

NeuralWorks Public NeuralWorks is a comprehensive & interactive toolkit for building, training, and running custom neural network models. Beginner friendly. Python ⭐ 1

NWCustom Public NeuralWorksCustom (NWCustom) Framework is a lightweight, customisable neural network toolkit inspired by PyTorch. Python

Homelab Public Always a hardware enthusiast since childhood, my custom homelab is where More coming soon...

Oguzhan Cagirir
OguzhanCOG · he/him

