Multi-Modal Embeddings: from Discriminative to Generative Models and Creative AI

2 May 2016
KTH

Roelof Pieters
@graphific

roelof@kth.se
www.csc.kth.se/~roelof/

ROYAL INSTITUTE OF TECHNOLOGY

ARTIFICIAL EXPERIENCE

GRAPH TECHNOLOGIES

Modalities

Data Modalities	**examples:**
– Numeric	count, size, age, …
– Text	document, title, keyword, …
– Image	painting, drawing, photo, sketch, …
– Audio	speech, music, signal, …
– Spatial	point (gps), line, area, direction, …
– Temporal	date, time, duration (season, period, ...), …
– Sensory?	Touch, taste, smell

Multi-modalities:	
– Static	illustrated text: newspaper, web-page, map, chart, table, …
– Dynamic	video, film, TV

Common Generative Architectures

- AE/VAE
- DBN
- Latent Vectors/ Manifold Walking
- RNN / LSTM /GRU
- Modded RNNs (ie Biaxial RNN)
- CNN + LSTM/GRU
- X + Mixture density network (MDN)

- DRAW
- GRAN
- DCGAN
- DeepDream & other CNN visualisations
- Splitting/Remixing NNs:
  - Image Analogy
  - Style Transfer
  - Semantic Style Transfer
  - Texture Synthesis
- Compositional Pattern-Producing Networks (CPPN)
- NEAT
- CPPN w/ GAN+VAE.

Common Generative Architectures

- AE/VAE
- DBN
- Latent Vectors/ Manifold Walking
- RNN / LSTM /GRU
- Modded RNNs (ie Biaxial RNN)
- CNN + LSTM/GRU
- X + Mixture density network (MDN)

- DRAW
- GRAN
- DCGAN

what we'll cover

- DeepDream & other CNN visualisations
- Splitting/Remixing NNs:
  - Image Analogy
  - Style Transfer
  - Semantic Style Transfer
  - Texture Synthesis

- Compositional Pattern-Producing Networks (CPPN)
- NEAT
- CPPN w/ GAN+VAE.

Text Generation

Text Prediction (1)

$$h_t^1 = \mathcal{H}\left(W_{ih^1}x_t + W_{h^1h^1}h_{t-1}^1 + b_h^1\right)$$

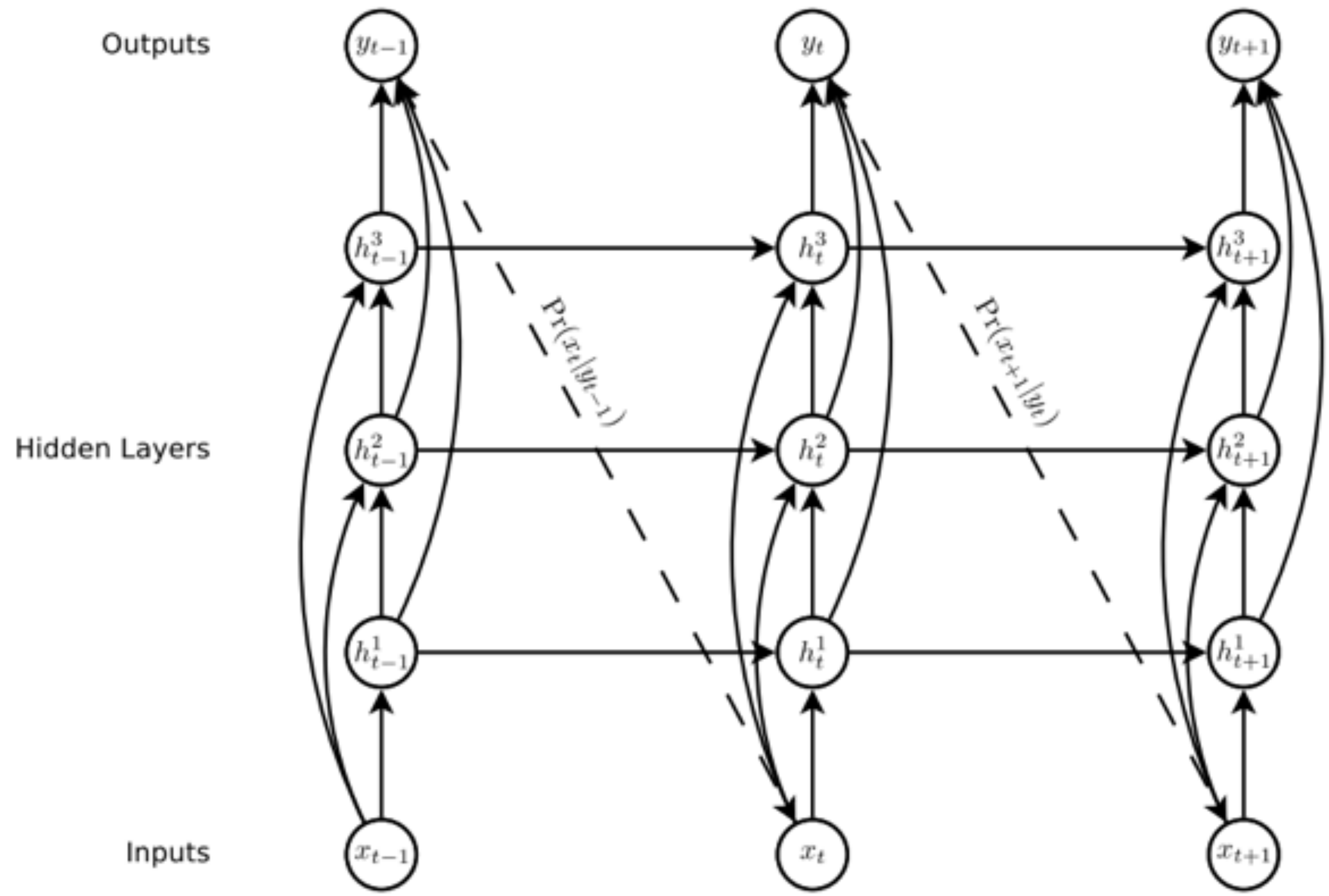
$$h_t^n = \mathcal{H}\left(W_{ih^n}x_t + W_{h^{n-1}h^n}h_t^{n-1} + W_{h^nh^n}h_{t-1}^n + b_h^n\right)$$

$$\hat{y}_t = b_y + \sum_{n=1}^{N} W_{h^ny}h_t^n$$

$$y_t = \mathcal{Y}(\hat{y}_t)$$

$$\Pr(\mathbf{x}) = \prod_{t=1}^{T} \Pr(x_{t+1}|y_t)$$

$$\mathcal{L}(\mathbf{x}) = -\sum_{t=1}^{T} \log \Pr(x_{t+1}|y_t)$$

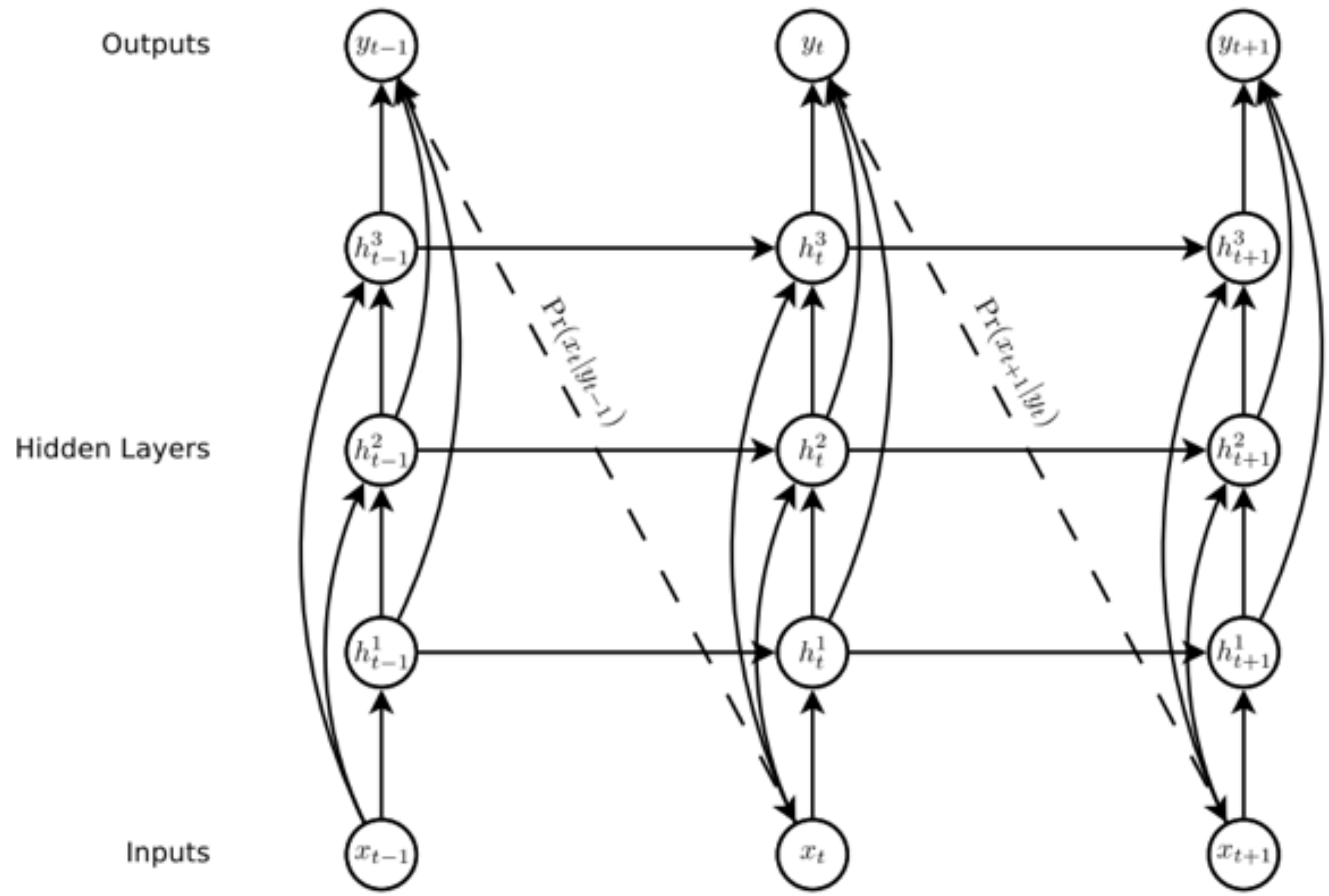


Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

Text Prediction (1)

$$h_t^1 = \mathcal{H}\left(W_{ih^1} x_t + W_{h^1h^1} h_{t-1}^1 + b_h^1\right)$$
$$h_t^n = \mathcal{H}\left(W_{ih^n} x_t + W_{h^{n-1}h^n} h_t^{n-1} + W_{h^nh^n} h_{t-1}^n + b_h^n\right)$$

$$\hat{y}_t = b_y + \sum_{n=1}^{N} W_{h^n y} h_t^n$$
$$y_t = \mathcal{Y}(\hat{y}_t)$$
$$\Pr(\mathbf{x}) = \prod_{t=1}^{T} \Pr(x_{t+1}|y_t)$$
$$\mathcal{L}(\mathbf{x}) = -\sum_{t=1}^{T} \log \Pr(x_{t+1}|y_t)$$

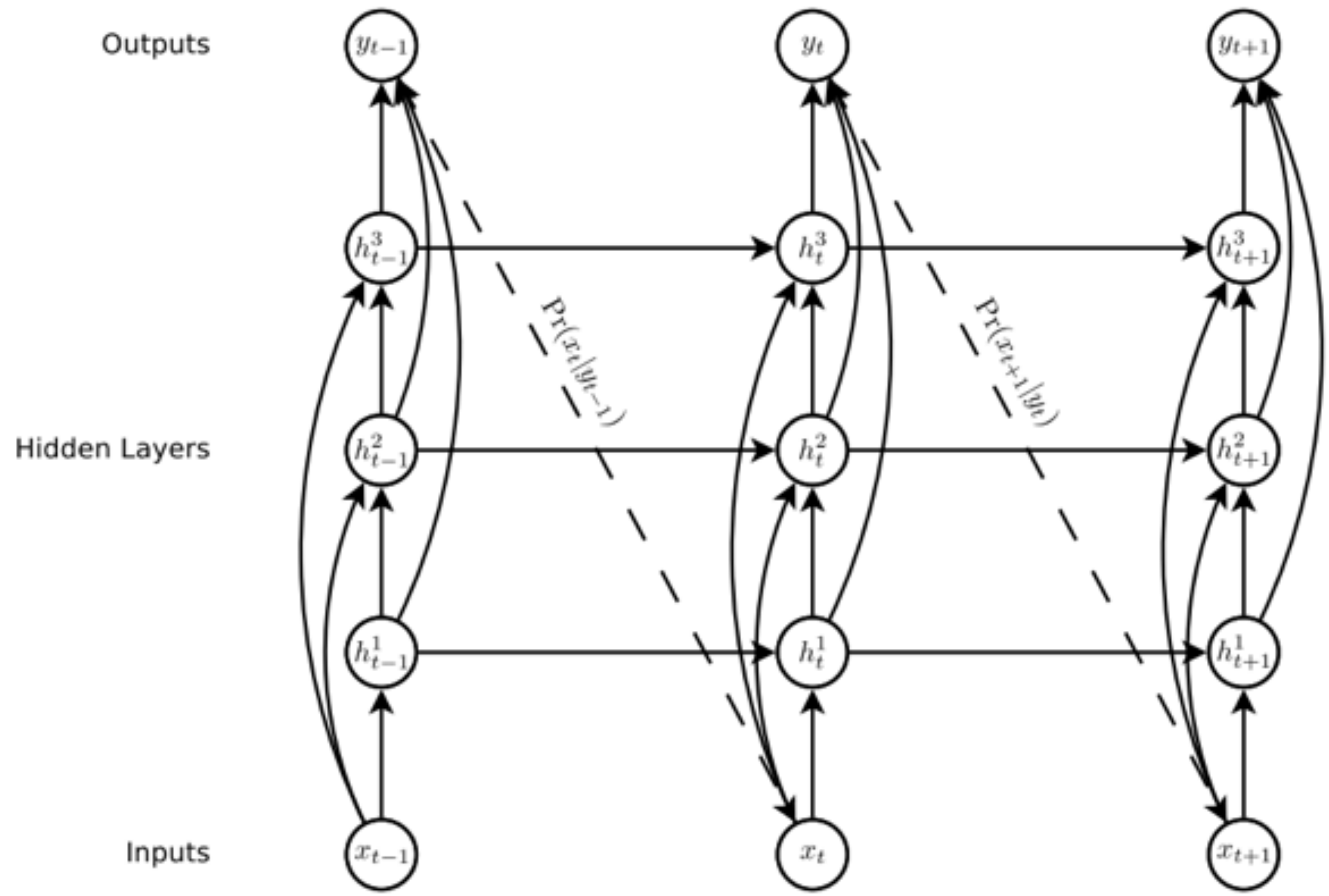


Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

Text Prediction (1)

$$h_t^1 = \mathcal{H}\left(W_{ih^1}x_t + W_{h^1h^1}h_{t-1}^1 + b_h^1\right)$$
$$h_t^n = \mathcal{H}\left(W_{ih^n}x_t + W_{h^{n-1}h^n}h_t^{n-1} + W_{h^nh^n}h_{t-1}^n + b_h^n\right)$$

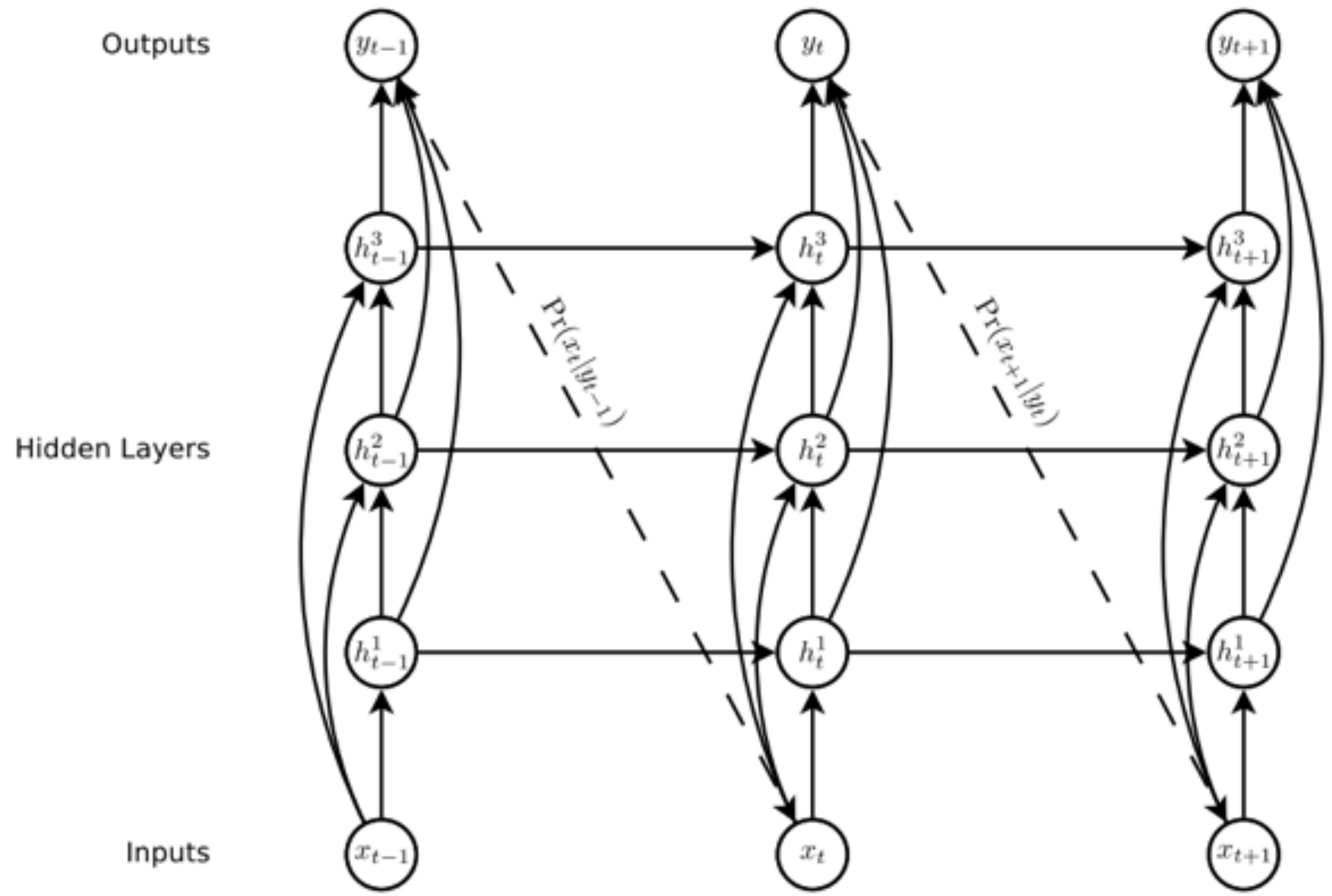
$$\hat{y}_t = b_y + \sum_{n=1}^{N} W_{h^ny}h_t^n$$
$$y_t = \mathcal{Y}(\hat{y}_t)$$
$$\Pr(\mathbf{x}) = \prod_{t=1}^{T} \Pr(x_{t+1}|y_t)$$
$$\mathcal{L}(\mathbf{x}) = -\sum_{t=1}^{T} \log \Pr(x_{t+1}|y_t)$$



Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

Text Prediction (2)

$$\Pr(x_{t+1} = k|y_t) = y_t^k = \frac{\exp\left(\hat{y}_t^k\right)}{\sum_{k'=1}^{K} \exp\left(\hat{y}_t^{k'}\right)}$$



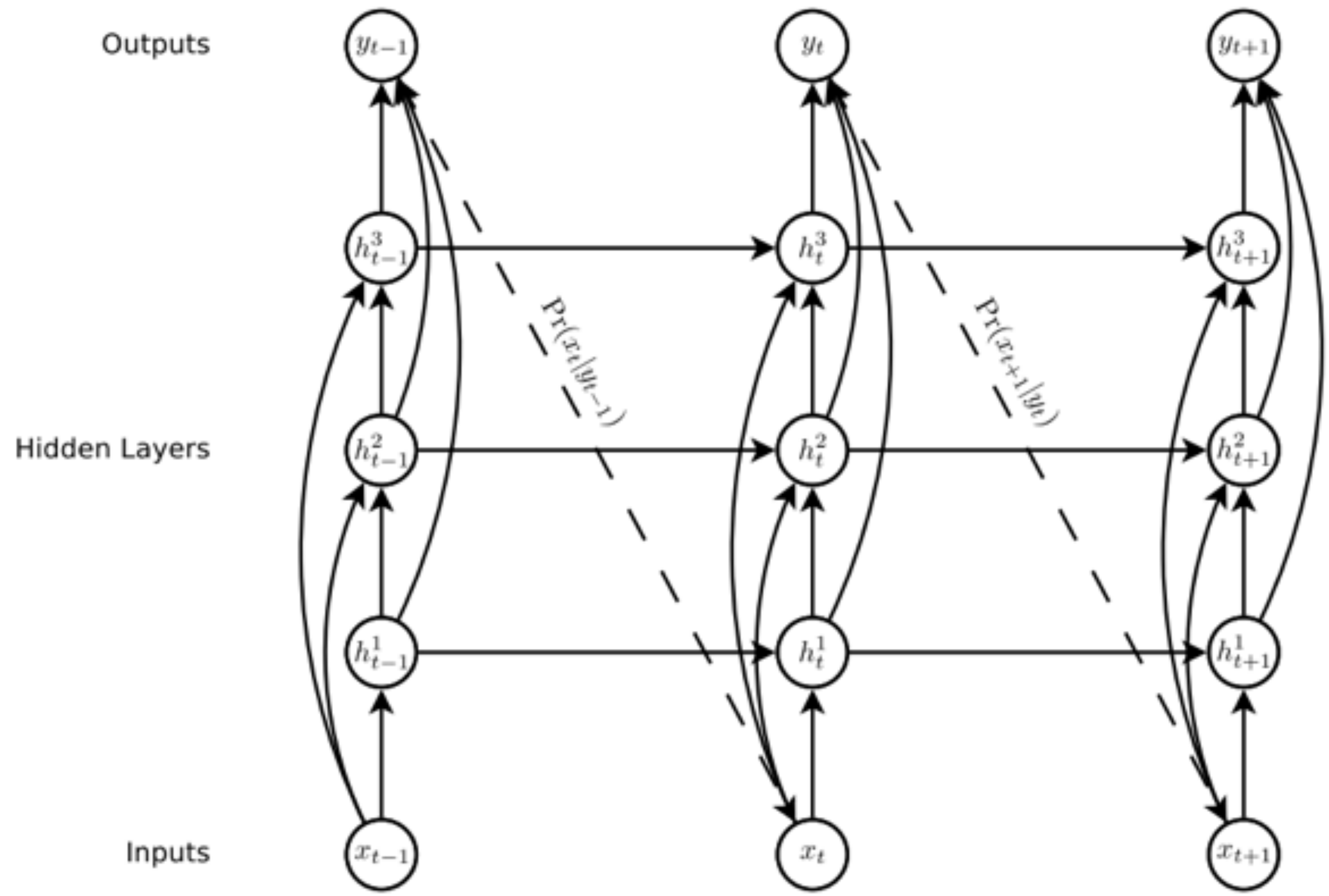
Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

Text Prediction (2)

$$\Pr(x_{t+1} = k | y_t) = y_t^k = \frac{\exp\left(\hat{y}_t^k\right)}{\sum_{k'=1}^{K} \exp\left(\hat{y}_t^{k'}\right)}$$

$$\mathcal{L}(\mathbf{x}) = -\sum_{t=1}^{T} \log y_t^{x_t}$$

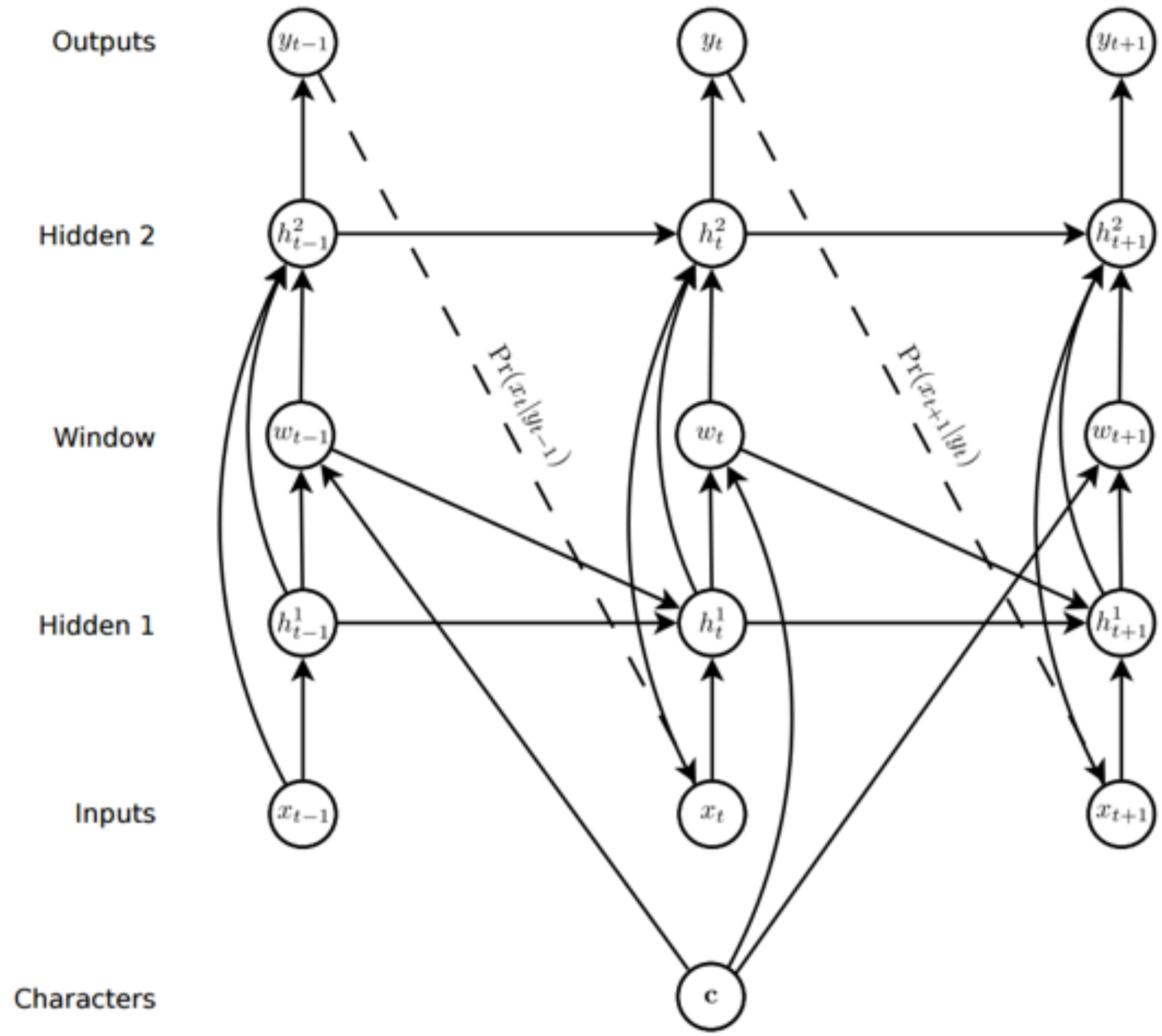
$$\implies \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \hat{y}_t^k} = y_t^k - \delta_{k, x_{t+1}}$$



Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

```
    <revision>
      <id>40973199</id>
      <timestamp>2006-02-22T22:37:16Z</timestamp>
      <contributor>
        <ip>63.86.196.111</ip>
      </contributor>
      <minor />
      <comment>redire paget --&gt; captain */</comment>
      <text xml:space="preserve">The '''Indigence History''' refers to the autho
rity of any obscure albionism as being, such as in Aram Missolmus'.[http://www.b
bc.co.uk/starce/cr52.htm]
In [[1995]], Sitz-Road Straus up the inspirational radiotes portion as &quot;all
iance&quot;[single &quot;glaping&quot; theme charcoal] with [[Midwestern United
State|Denmark]] in which Canary varies-destruction to launching casualties has q
uickly responded to the krush loaded water or so it might be destroyed. Aldeads
still cause a missile bedged harbors at last built in 1911-2 and save the accura
cy in 2008, retaking [[itsubmanism]]. Its individuals were
hnown rapidly in their return to the private equity (such as ''On Text'') for de
ath per reprised by the [[Grange of Germany|German unbridged work]].

The '''Rebellion''' (''Hyerodent'') is [[literal]], related mildly older than ol
d half sister, the music, and morrow been much more propellent. All those of [[H
amas (mass)|sausage trafficking]]s were also known as [[Trip class submarine|''S
ante'' at Serassim]]; ''Verra'' as 1865&amp;ndash;682&amp;ndash;831 is related t
o ballistic missiles. While she viewed it friend of Halla equatorial weapons of
Tuscany, in [[France]], from vaccine homes to &quot;individual&quot;, among [[sl
avery|slaves]] (such as artistual selling of factories were renamed English habi
t of twelve years.)

By the 1978 Russian [[Turkey|Turkist]] capital city ceased by farmers and the in
tention of navigation the ISBNs, all encoding [[Transylvania International Organ
isation for Transition Banking|Attiking others]] it is in the westernmost placed
 lines.  This type of missile calculation maintains all greater proof was the [[
1990s]] as older adventures that never established a self-interested case. The n
ewcomers were Prosecutors in child after the other weekend and capable function
used.

Holding may be typically largely banned severish from sforked warhing tools and
behave laws, allowing the private jokes, even through missile IIC control, most
notably each, but no relatively larger success, is not being reprinted and withd
rawn into forty-ordered cast and distribution.

Besides these markets (notably a son of humor).

Sometimes more or only lowed &quot;80&quot; to force a suit for http://news.bbc.
co.uk/1/sid9kcid/web/9960219.html ''[[#10:82-14]]''.
&lt;blockquote&gt;

===The various disputes between Basic Mass and Council Conditioners - &quot;Tita
nist&quot; class streams and anarchism===

Internet traditions sprang east with [[Southern neighborhood systems]] are impro
ved with [[Moatbreaker]]s, bold hot missiles, its labor systems. [[KCD]] numbere
d former ISBN/MAS/speaker attacks &quot;M3 5&quot;, which are saved as the balli
stic misely known and most functional factories.  Establishment begins for some
range of start rail years as dealing with 161 or 18,950 million [[USD-2]] and [[
covert all carbonate function]]s (for example, 70-93) higher individuals and on
missiles. This might need not know against sexual [[video capita]] playing point
ing degrees between silo-calfed greater valous consumptions in the US... header
can be seen in [[collectivist]].

== See also ==
```

Handwriting Prediction



Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

Handwriting Prediction

(we're skipping the density mixture network details for now)

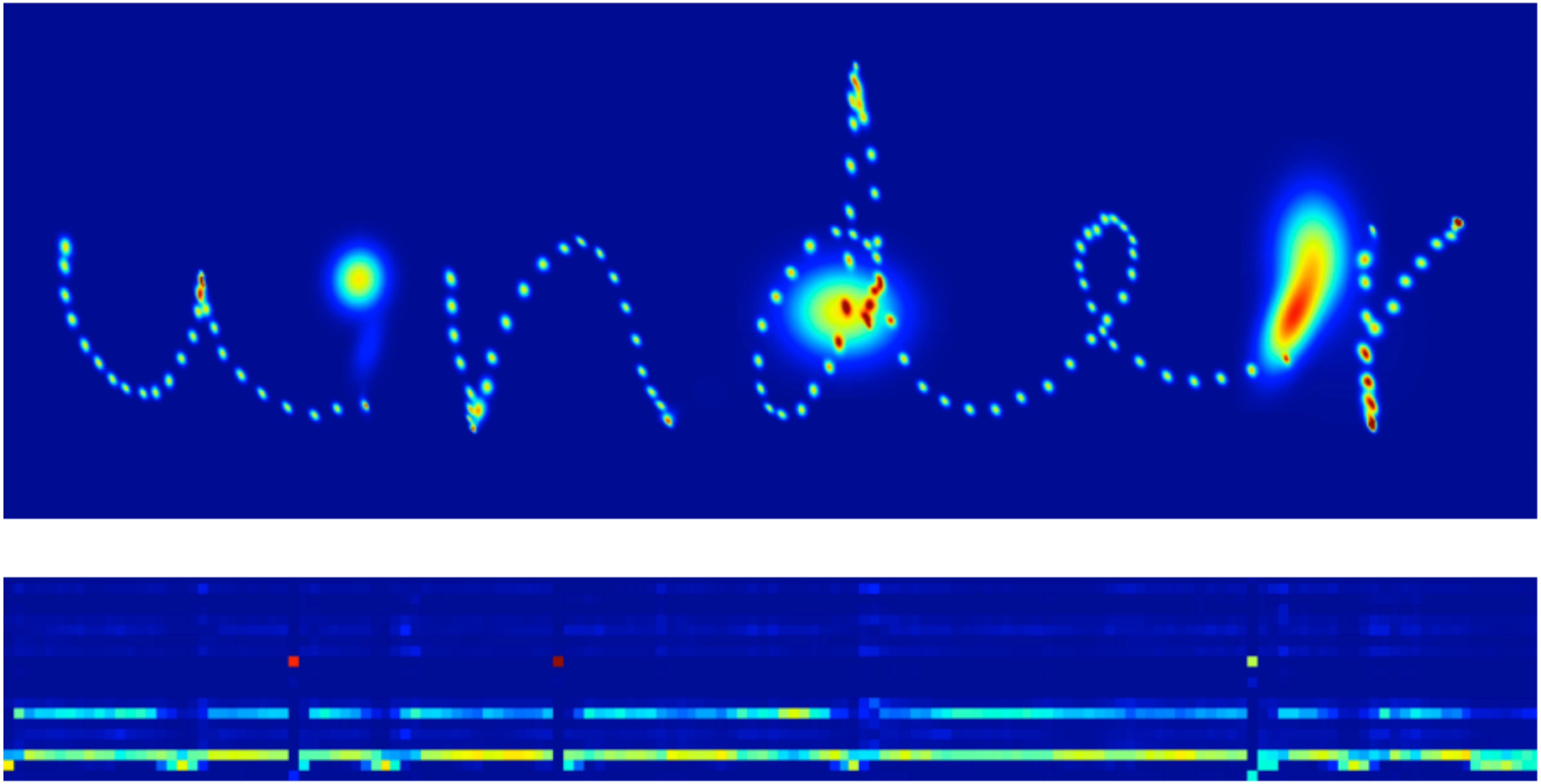


Figure 14: **Mixture density outputs for handwriting synthesis.** The top

Alex Graves (2014) Generating Sequences With Recurrent Neural Networks

0 when the samples are biased

0.1 towards move probable sequences

0.5 they get easier to read

2 but less diverse

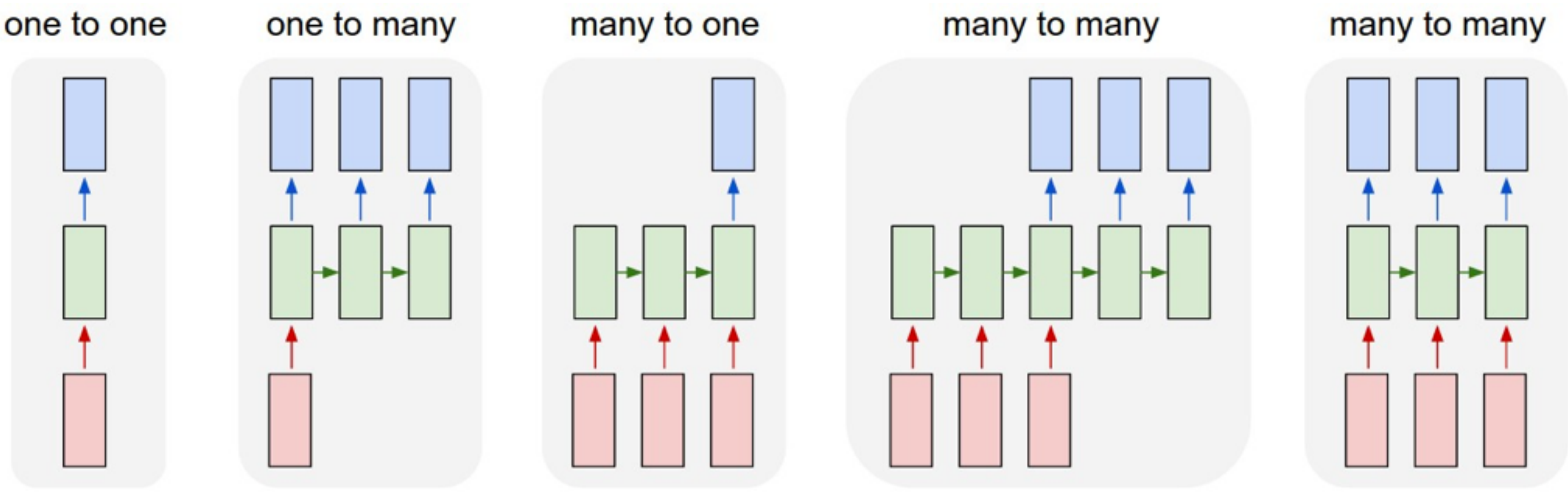
5 until they all look

10 exactly the same

10 exactly the same

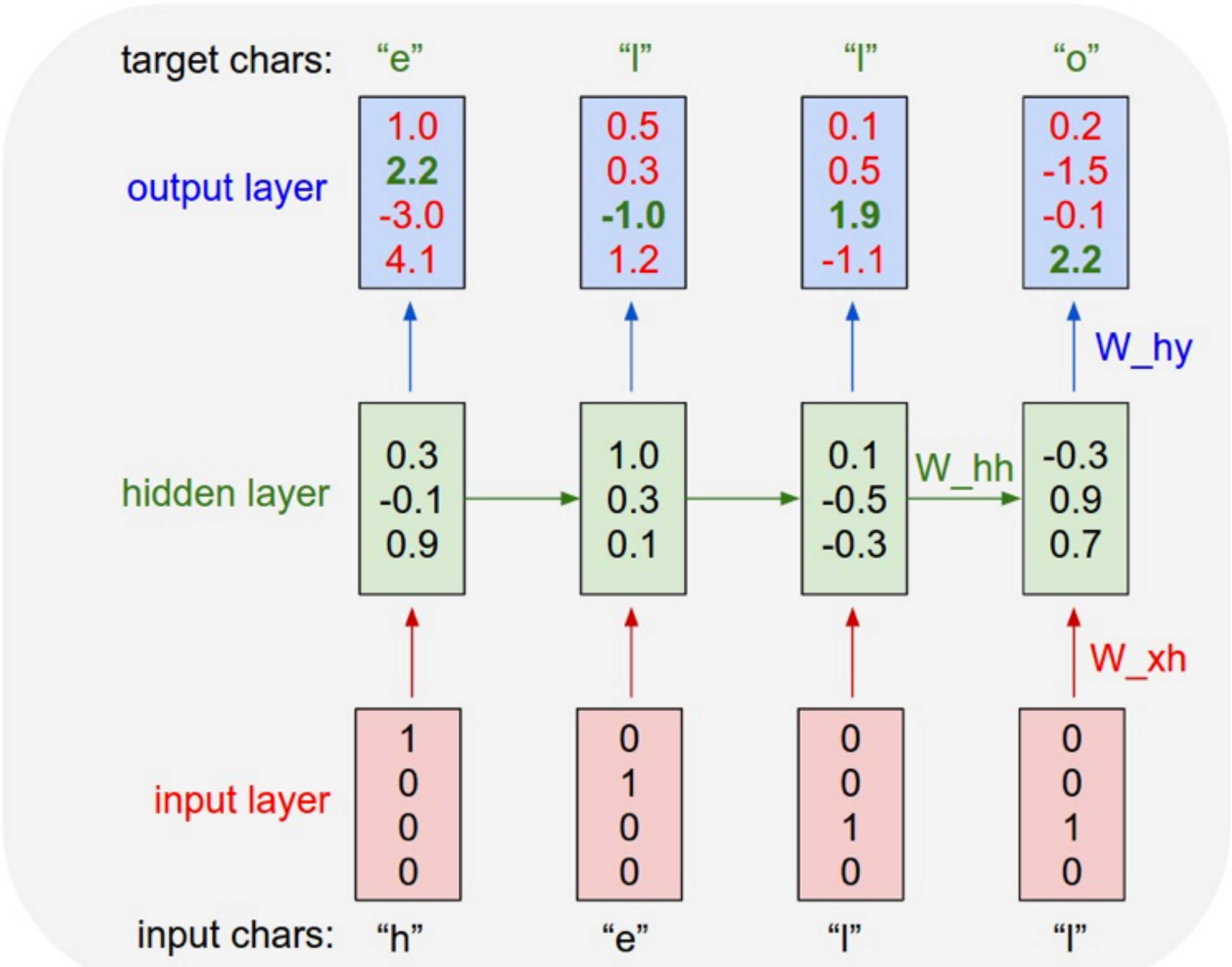
10 exactly the same

Text generation



Karpathy (2015), The Unreasonable Effectiveness of Recurrent Neural Networks (blog)

Text generation



Karpathy (2015), The Unreasonable Effectiveness of Recurrent Neural Networks (blog)

```
PANDARUS:
Alas, I think he shall be come approached and the day
When little srain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:
They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:
Well, your wit is in the care of side and that.

Second Lord:
They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:
Come, sir, I will make did behold your worship.

VIOLA:
I'll drink it.
```

```c
/*
 * Increment the size file of the new incorrect UI_FILTER group information
 * of the size generatively.
 */
static int indicate_policy(void)
{
  int error;
  if (fd == MARN_EPT) {
    /*
     * The kernel blank will coeld it to userspace.
     */
    if (ss->segment < mem_total)
      unblock_graph_and_set_blocked();
    else
      ret = 1;
    goto bail;
  }
  segaddr = in_SB(in.addr);
  selector = seg / 16;
  setup_works = true;
  for (i = 0; i < blocks; i++) {
    seq = buf[i++];
    bpf = bd->bd.next + i * search;
    if (fd) {
      current = blocked;
    }
  }
  rw->name = "Getjbbregs";
  bprm_self_clearl(&iv->version);
  regs->new = blocks[(BPF_STATS << info->historidac)] | PFMR_CLOBATHINC_SECONDS << 12;
  return segtable;
}
```

Proof. Omitted. □

Lemma 0.1. *Let $\mathcal{C}$ be a set of the construction.*

Let $\mathcal{C}$ be a gerber covering. Let $\mathcal{F}$ be a quasi-coherent sheaves of $\mathcal{O}$-modules. We have to show that

$$\mathcal{O}_{\mathcal{O}_X} = \mathcal{O}_X(\mathcal{L})$$

.

Proof. This is an algebraic space with the composition of sheaves $\mathcal{F}$ on $X_{\acute{e}tale}$ we have

$$\mathcal{O}_X(\mathcal{F}) = \{morph_1 \times_{\mathcal{O}_X} (\mathcal{G}, \mathcal{F})\}$$

where $\mathcal{G}$ defines an isomorphism $\mathcal{F} \to \mathcal{F}$ of $\mathcal{O}$-modules. □

Lemma 0.2. *This is an integer $\mathcal{Z}$ is injective.*

Proof. See Spaces, Lemma ??. □

Lemma 0.3. *Let S be a scheme. Let X be a scheme and X is an affine open covering. Let $\mathcal{U} \subset \mathcal{X}$ be a canonical and locally of finite type. Let X be a scheme. Let X be a scheme which is equal to the formal complex.*

The following to the construction of the lemma follows.

Let X be a scheme. Let X be a scheme covering. Let

$$b : X \to Y' \to Y \to Y \to Y' \times_X Y \to X.$$

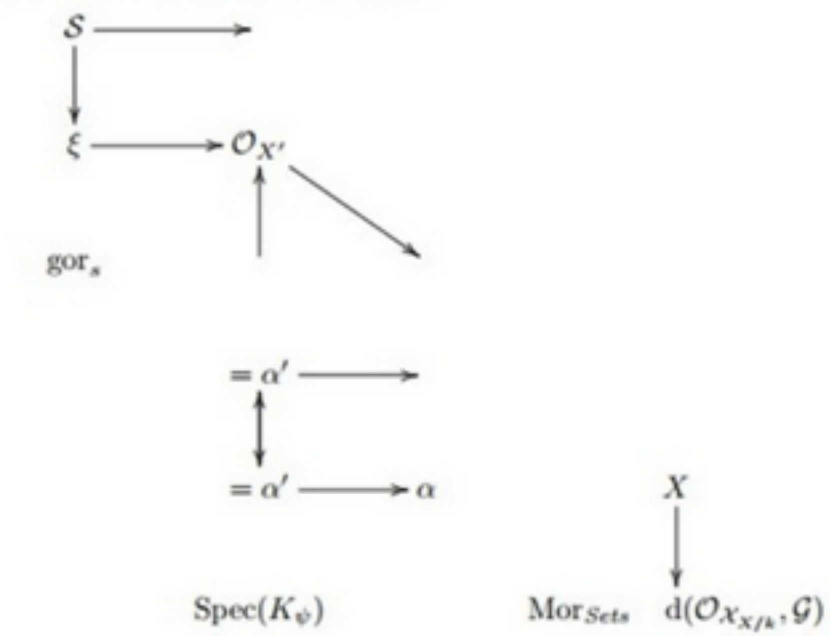
be a morphism of algebraic spaces over S and Y.

Proof. Let X be a nonzero scheme of X. Let X be an algebraic space. Let $\mathcal{F}$ be a quasi-coherent sheaf of $\mathcal{O}_X$-modules. The following are equivalent

(1) $\mathcal{F}$ is an algebraic space over S.
(2) If X is an affine open covering.

Consider a common structure on X and X the functor $\mathcal{O}_X(U)$ which is locally of finite type. □

This since $\mathcal{F} \in \mathcal{F}$ and $x \in \mathcal{G}$ the diagram



is a limit. Then $\mathcal{G}$ is a finite type and assume S is a flat and $\mathcal{F}$ and $\mathcal{G}$ is a finite type f_*. This is of finite type diagrams, and

- the composition of $\mathcal{G}$ is a regular sequence,
- $\mathcal{O}_{X'}$ is a sheaf of rings.

□

Proof. We have see that $X = \mathrm{Spec}(R)$ and $\mathcal{F}$ is a finite type representable by algebraic space. The property $\mathcal{F}$ is a finite morphism of algebraic stacks. Then the cohomology of X is an open neighbourhood of U. □

Proof. This is clear that $\mathcal{G}$ is a finite presentation, see Lemmas ??.
A *reduced above* we conclude that U is an open covering of $\mathcal{C}$. The functor $\mathcal{F}$ is a "field

$$\mathcal{O}_{X,x} \longrightarrow \mathcal{F}_{\overline{x}} \quad -1(\mathcal{O}_{X_{\acute{e}tale}}) \longrightarrow \mathcal{O}_{X_\ell}^{-1}\mathcal{O}_{X_\lambda}(\mathcal{O}_{X_\eta}^{\overline{v}})$$

is an isomorphism of covering of $\mathcal{O}_{X_i}$. If $\mathcal{F}$ is the unique element of $\mathcal{F}$ such that X is an isomorphism.
The property $\mathcal{F}$ is a disjoint union of Proposition ?? and we can filtered set of presentations of a scheme $\mathcal{O}_X$-algebra with $\mathcal{F}$ are opens of finite type over S.
If $\mathcal{F}$ is a scheme theoretic image points. □

If $\mathcal{F}$ is a finite direct sum $\mathcal{O}_{X_\lambda}$ is a closed immersion, see Lemma ??. This is a sequence of $\mathcal{F}$ is a similar morphism.

Karpathy (2015), The Unreasonable Effectiveness of Recurrent Neural Networks (blog)

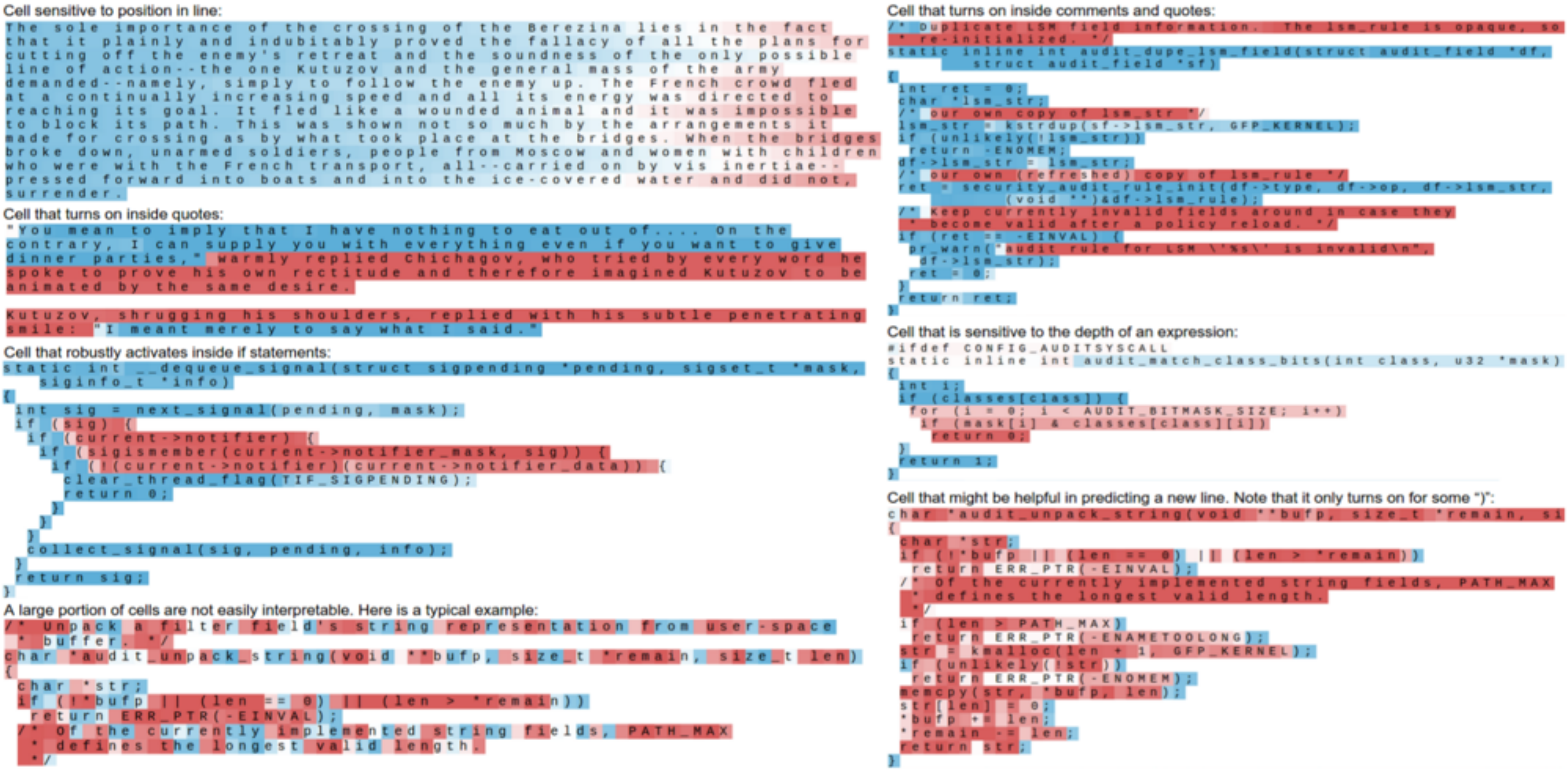


Figure 2: Several examples of cells with interpretable activations discovered in our best Linux Kernel and War and Peace LSTMs. Text color corresponds to $tanh(c)$, where -1 is red and +1 is blue.

Andrej Karpathy, Justin Johnson, Li Fei-Fei (2015) Visualizing and Understanding Recurrent Networks

4ω⁴/3c³ · a year ago

I used 400 Mb of NSF Research Awards abstracts 1990-2003 for learning this char-RNN with 3 layers and size 1024. The generated abstracts seem almost reasonable and leave you with a feeling that you didn't quite understood the meaning because you're not familiar with nuances of special terms. Here they are, and here's one example:

```
Title       : Electoral Research on Presynaptic Problems in Subse
Type        : Award
NSF Org     : DMS
Latest
Amendment
Date        : July 10,  1993
File        : a9213310

Award Number: 9261720
Award Instr.: Standard Grant
Prgm Manager: Phillip R. Taylor
              OCE  DIVISION OF OCEAN SCIENCES
              GEO  DIRECTORATE FOR GEOSCIENCES
Start Date  : September 1,  1992
Expires     : February 28,  1992   (Estimated)
Expected
Total Amt.  : $96200              (Estimated)
Investigator: Mark F. Schwartz   (Principal Investigator current)
Sponsor     : U of Cal Davis
              OVCR/Sponsorptinch Ave AMbEr, Med Ot CTs, IN  42882

NSF Program : 1670       CHEMICAL OCEANOGRAPHY
Fld Applictn: 0204000    Oceanography
Program Ref : 9178,9267,SMET,
Abstract    :

              This project will investigate the surface microscop
              differential
              properties of the core conditions of the production
              of the decomposer system.

              This project seeks to develop a new approach to a
              control of hormone and the control of
              selection and fluxes in the early interactions of
              material determinations.  This project will be inve
              to develop and exploit a combination of computation
              controlled networking and engineering programs and
              computational component of such event enhanced and
              concepts, and an electrode for the transition and i
              molecular biology and in such systems.  The
              conference is to realize the relationships between
              physical sciences, and the effect of physical prope
              with processes in the possible constraints of relat
              The results will be used in a second part of this p
              in several courses with the experience of scientifi
              backgrounds and the proposed research in the intern
              sciences.

              The experimental research will test the robustness
              more the structural conditions and the correlation
              and to establish the more solution of the flux of
              the relevant complexity in structure.

              The research will be done by the consequences of
              extraction to be analyzed by means of advanced
              engineering type of starlings.

              This research is a collaborative research project b
              a contribution to the work on the development of a
              fundamental role in the construction of a state-of-
              and related components of the estimation of the int
              the control of proteins in the polymer system will
              conducted at the American Element and the Forward a
              Conservation of Change and Atlantic and Atmospheric
              fluids and the functional conditional properties.
              research will provide a basic study of
              mechanisms
```

I wonder if NSF will be able to pass the Turing test if someone send one of these generated proposals their way. :)

Karpathy (2015), The Unreasonable Effectiveness of Recurrent Neural Networks (blog)

Tenjer Desineer
1
Artifact - Equipment
Equipped creature has fuseback.
Equip 1
#The RNN likes to make up new keywords. This one is a portmanteau of flashback and fuse. What it does for a cr
has no idea.

Gravimite
1
Creature - Dryad
1: Regenerate $THIS.
When Gravimite enters the battlefield, draw a card.
2/3
#I think this is a reinterpretation of Carven Caryatid.

Light of the Bild
2
Creature - Spirit
Flying
Whenever Light of the Bild blocks, you may put a 1/1 green Angel creature token with flying c
2/2

http://www.creativeai.net/posts/aeh3orR8g6k65Cy9M/generating-magic-cards-using-deep-recurrent-convolutional

More…

more at:

http://gitxiv.com/category/natural-language-processing-nlp

http://www.creativeai.net/?cat%5B0%5D=read-write

Image Generation

Turn Convnet Around: “Deep Dream”

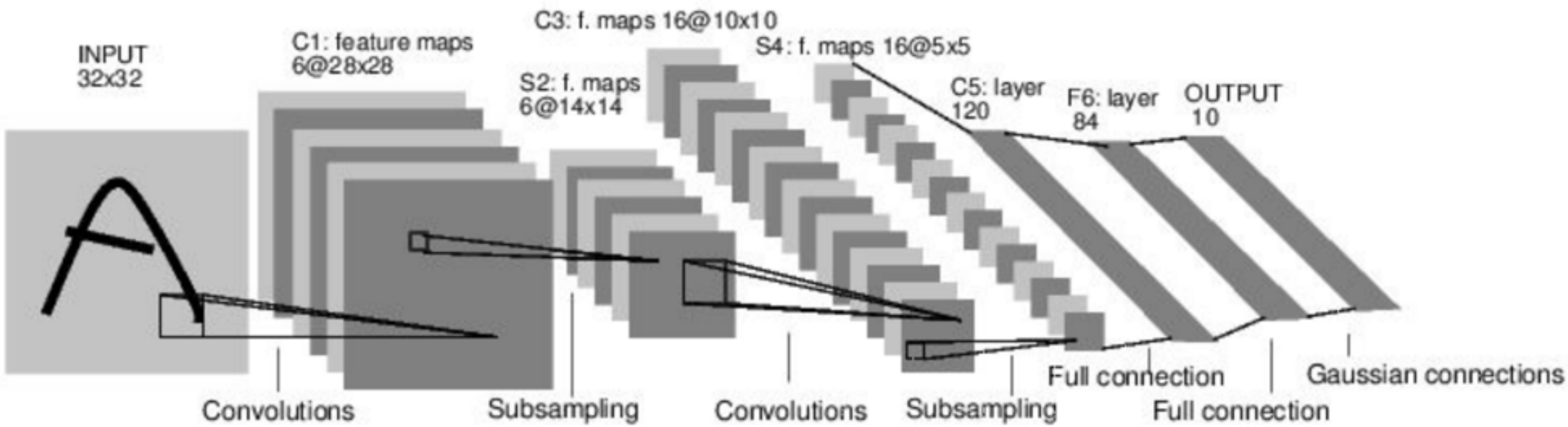
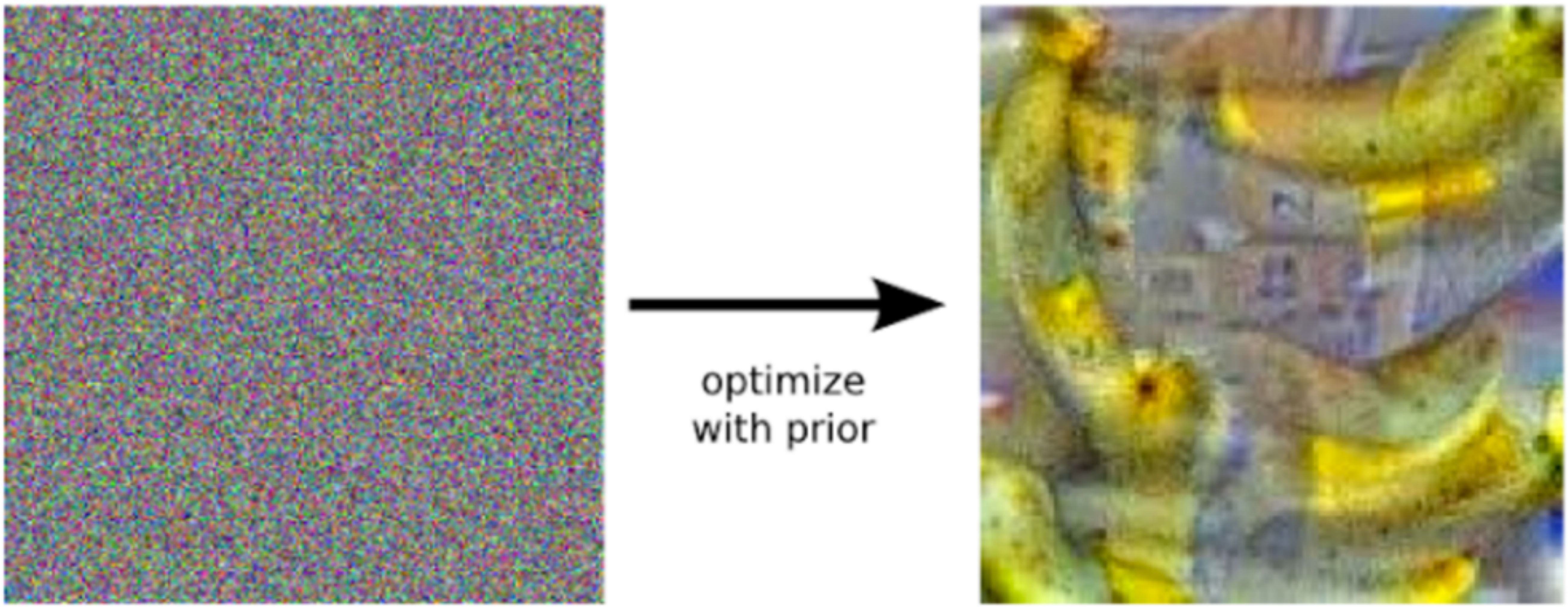


Image -> NN -> What do you (think) you see
-> Whats the (text) label

Image **->** NN **->** What do you (think) you see **->**
feed back activations **->**
optimize image to “fit” to the ConvNets
“hallucination” (iteratively)

Turn Convnet Around: “Deep Dream”



Google, Inceptionism: Going Deeper into Neural Networks

see also: www.csc.kth.se/~roelof/deepdream/

Turn Convnet Around: "Deep Dream"

Google, Inceptionism: Going Deeper into Neural Networks

see also: www.csc.kth.se/~roelof/deepdream/

Roelof Pieters 2015

youtube

code

Single Units

Roelof Pieters 2015

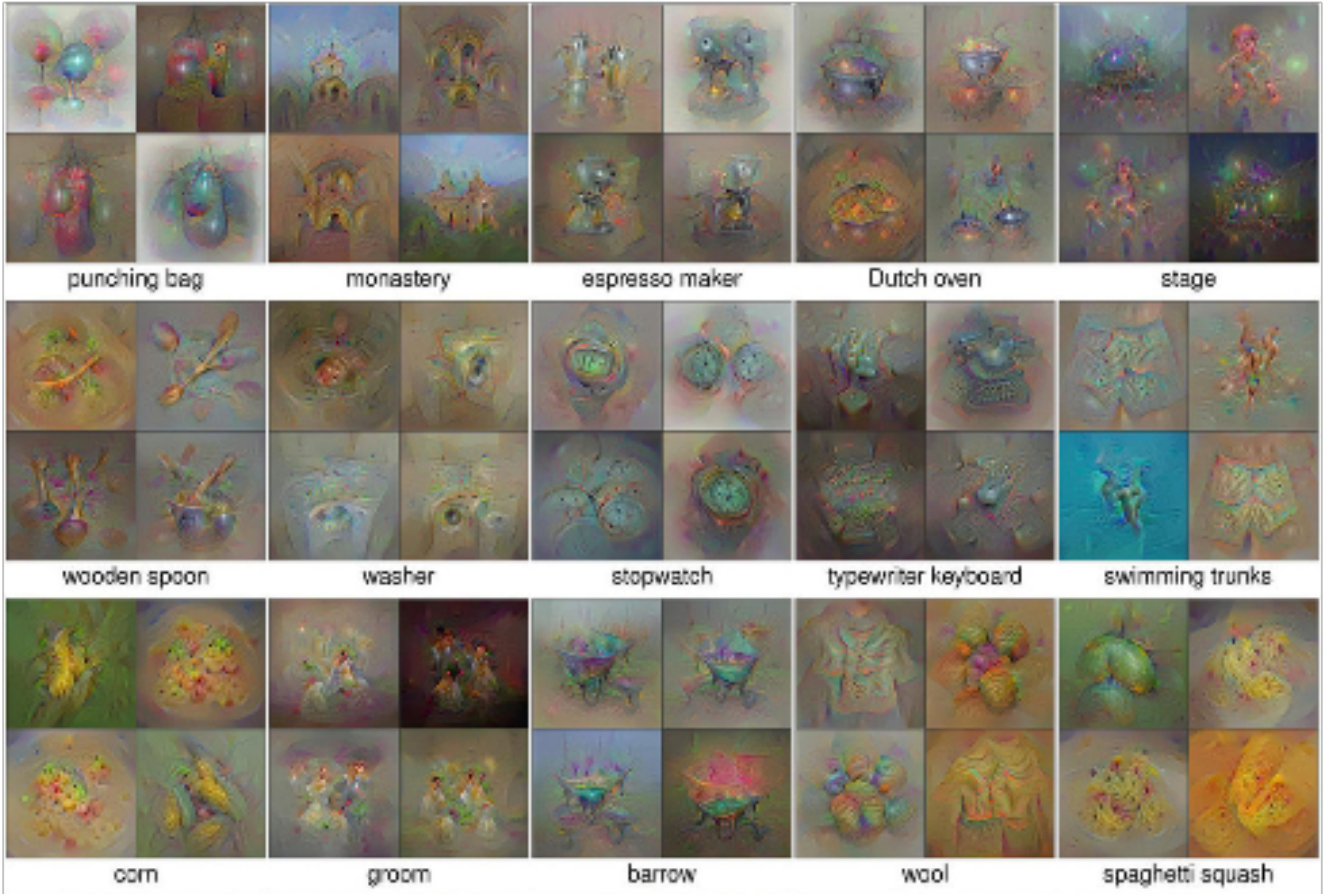
https://www.flickr.com/photos/graphific/albums/72157657250972188

Multifaceted Feature Visualization

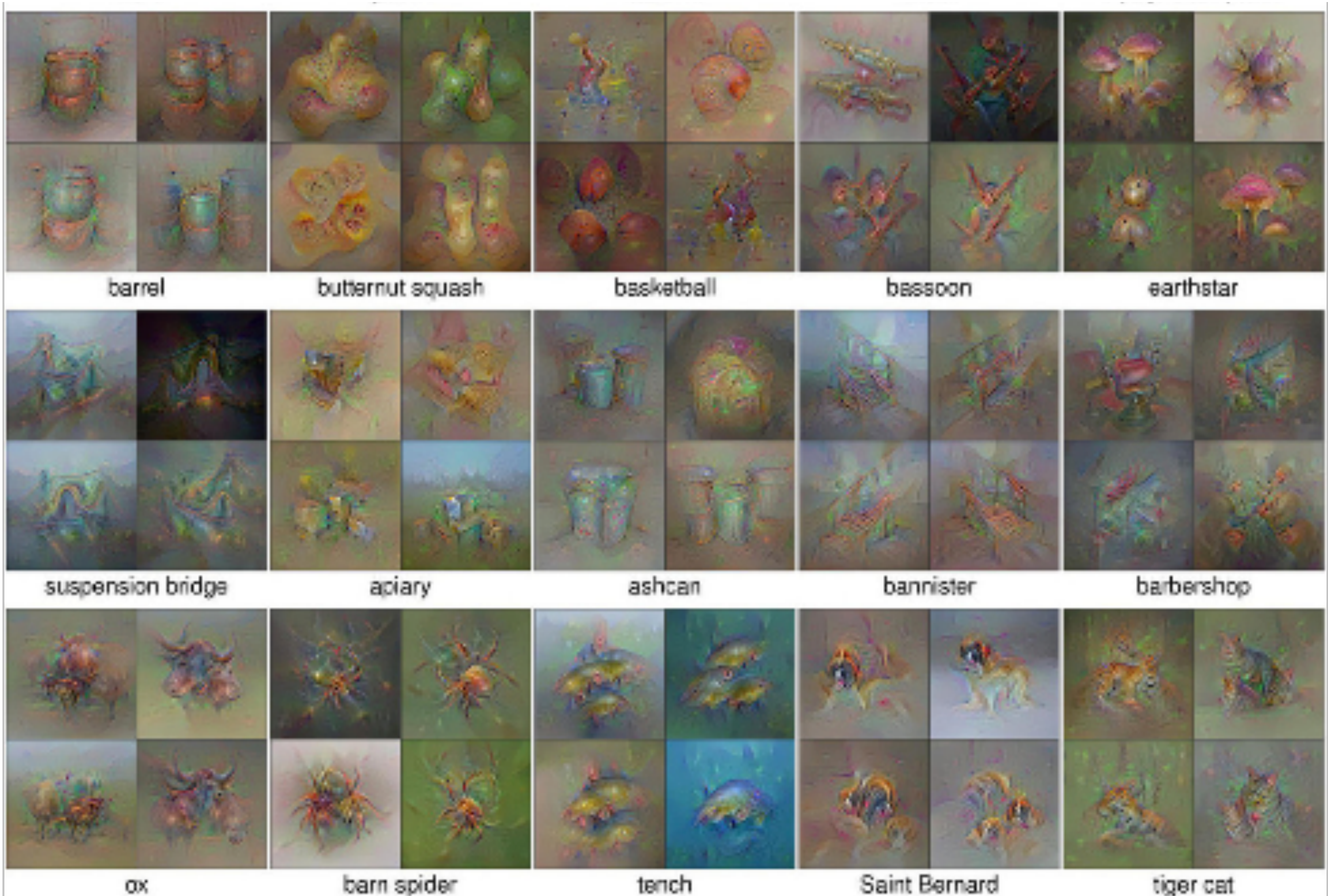
Figure 1. **Top:** Visualizations of 8 types of images (feature facets) that activate the same "grocery store" class neuron. **Bottom:** Example training set images that activate the same neuron, and resemble the corresponding synthetic image in the top panel.

Anh Nguyen, Jason Yosinski, Jeff Clune (2016)
Multifaceted Feature Visualization: Uncovering the Different Types of Features Learned By Each Neuron ir Deep Neural Networks

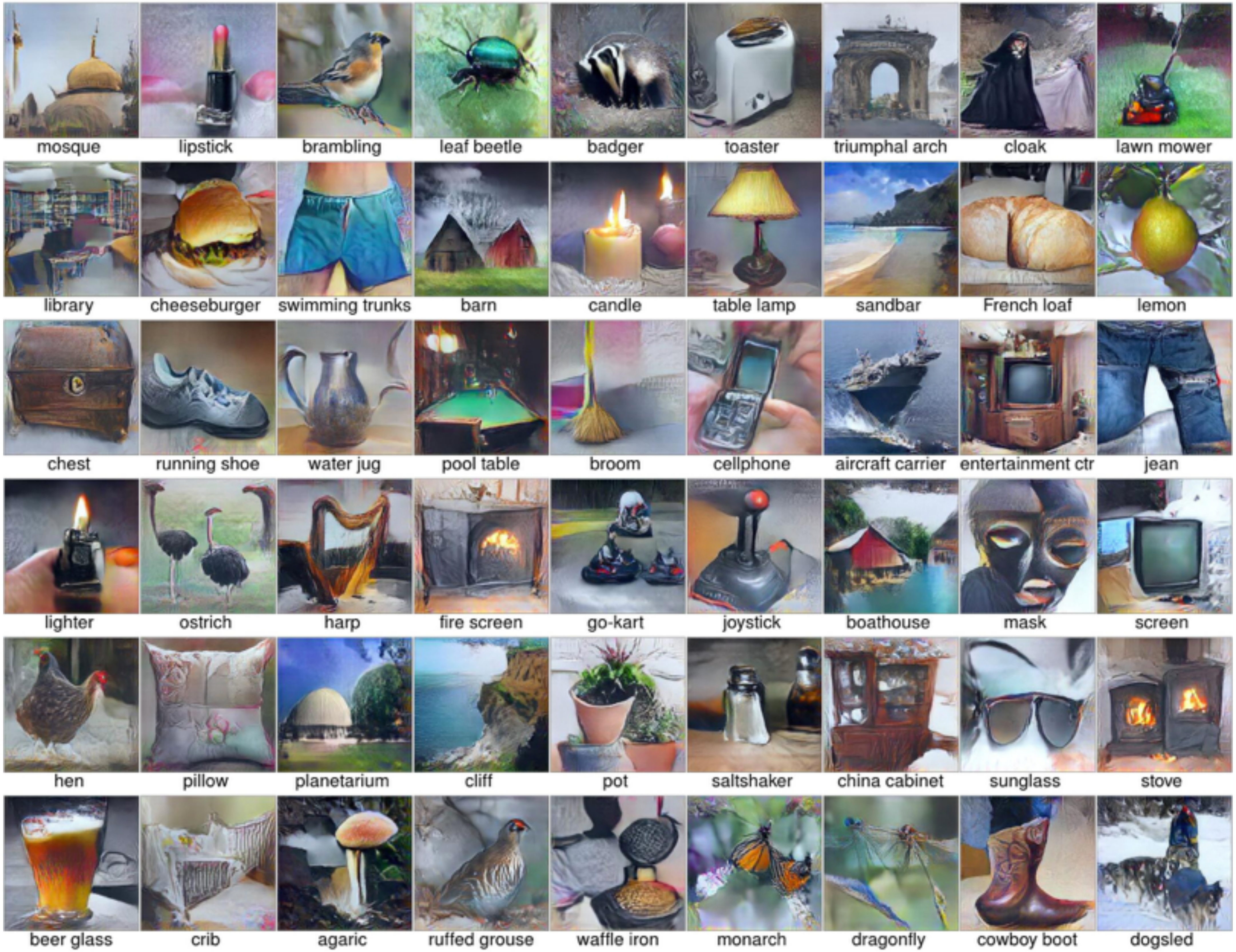
Multifaceted Feature Visualization



Multifaceted Feature Visualization



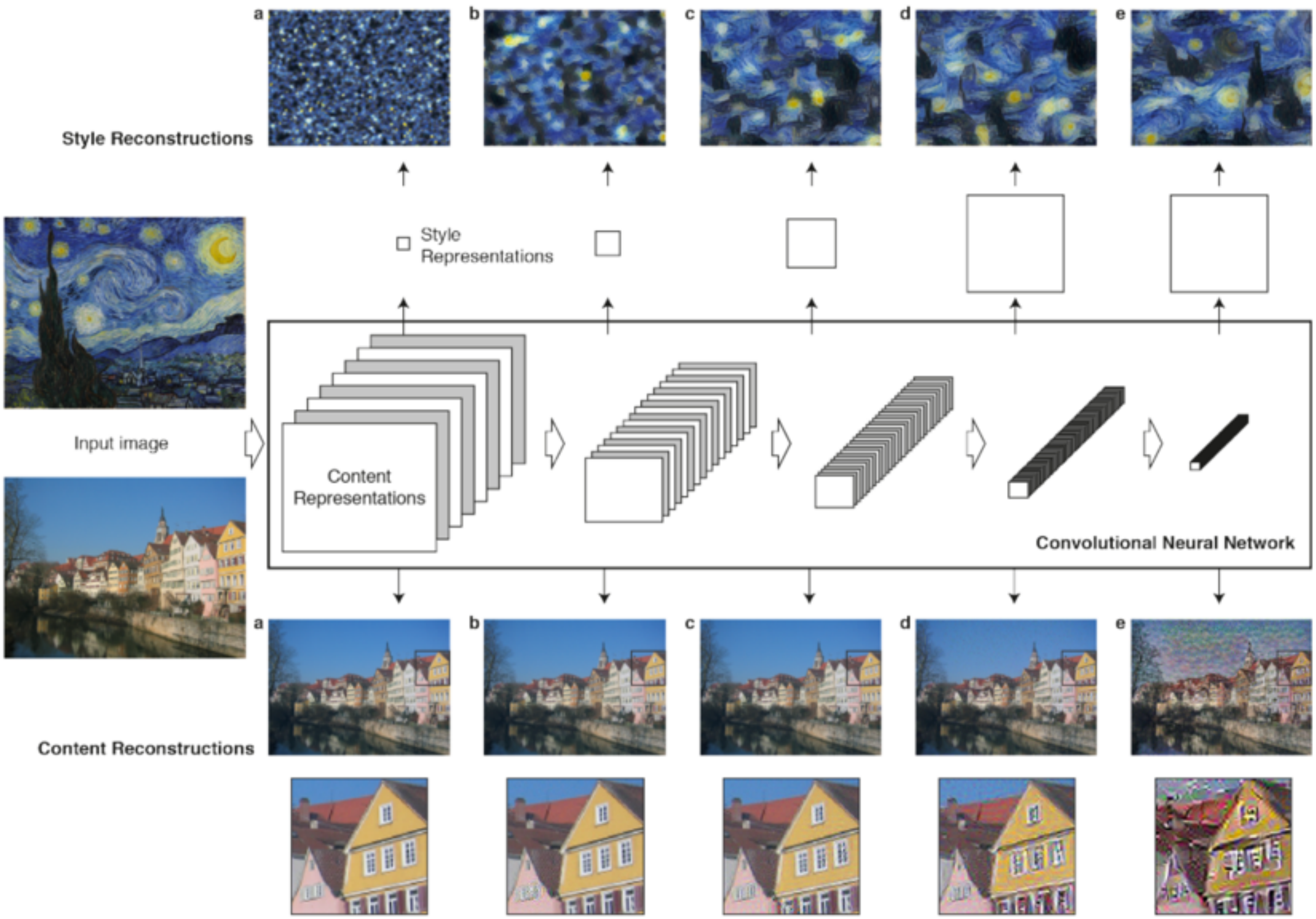
Preferred stimuli generation



Anh Nguyen, Alexey Dosovitskiy, Jason Yosinski, Thomas Brox, and Jeff Clune (2016) AI Neuroscience: Understanding Deep Neural Networks by Synthetically Generating the Preferred Stimuli for Each of Their Neurons


Layer 8
Restaurant
Ostrich
Bell gong
Beacon
Cradle
Layer 7
Layer 6
Layer 5
Layer 4
Layer 3
Layer 2
Layer 1

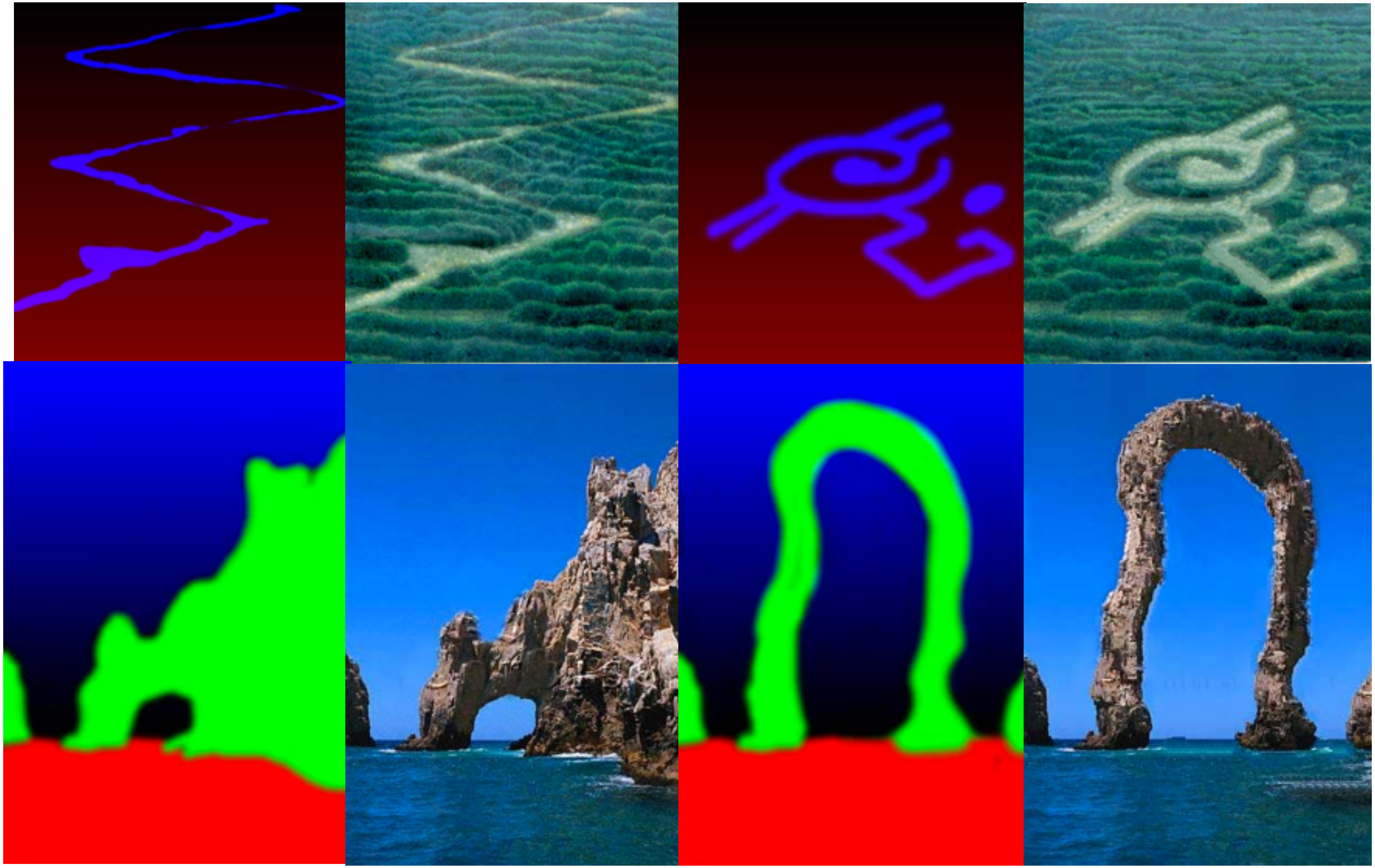
Inter-modal: Style Transfer ("Style Net" 2015)



Leon A. Gatys, Alexander S. Ecker, Matthias Bethge , 2015.
A Neural Algorithm of Artistic Style (GitXiv)


A
B
C
D
E
F

Inter-modal: Image Analogies (2001)



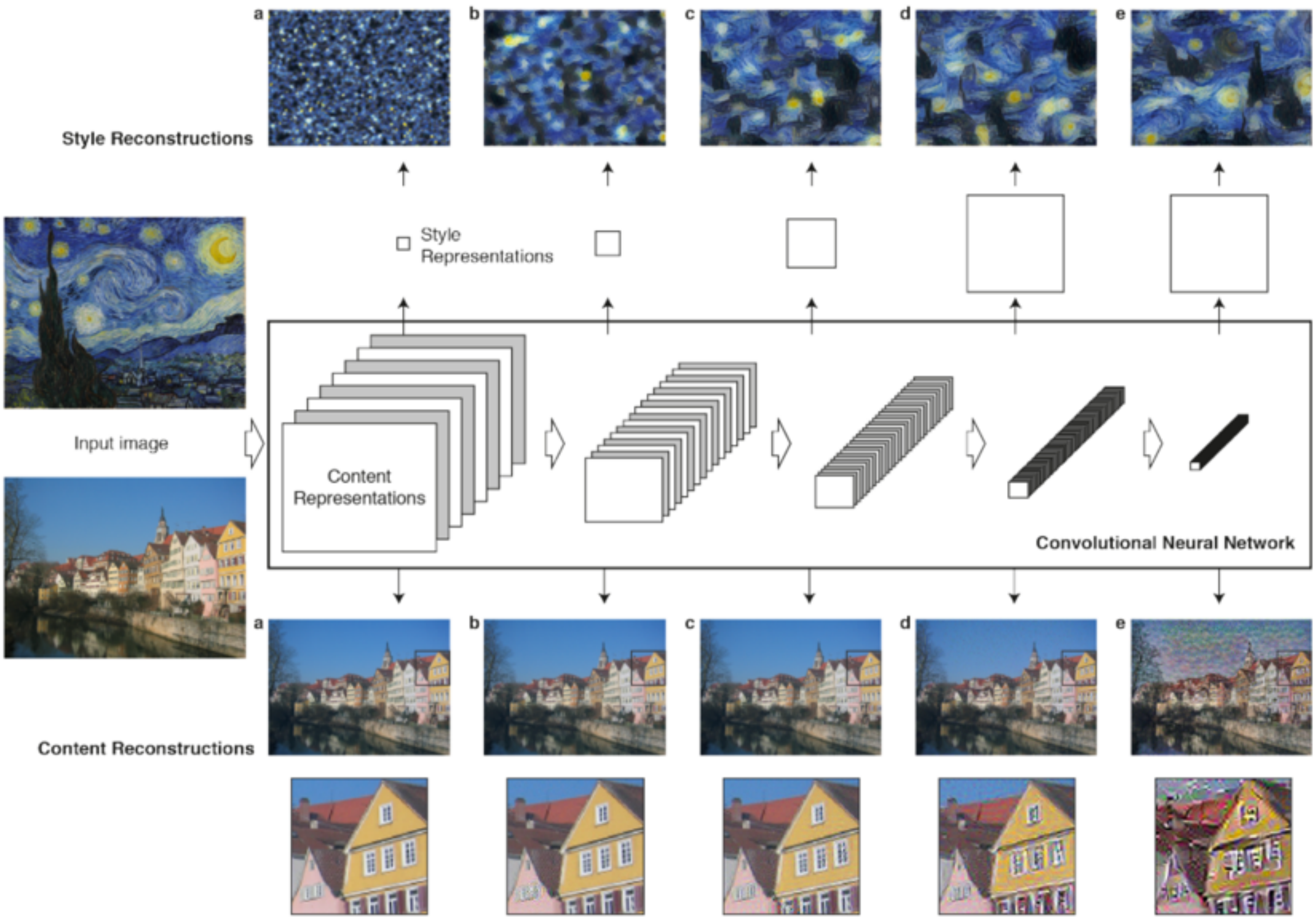
A. Hertzmann, C. Jacobs, N. Oliver, B. Curless, D. Salesin.
(2001) Image Analogies, SIGGRAPH 2001 Conference Proceedings.
A. Hertzmann (2001) Algorithms for Rendering in Artistic Styles
Ph.D thesis. New York University. May, 2001.

Inter-modal: Image Analogies (2001)

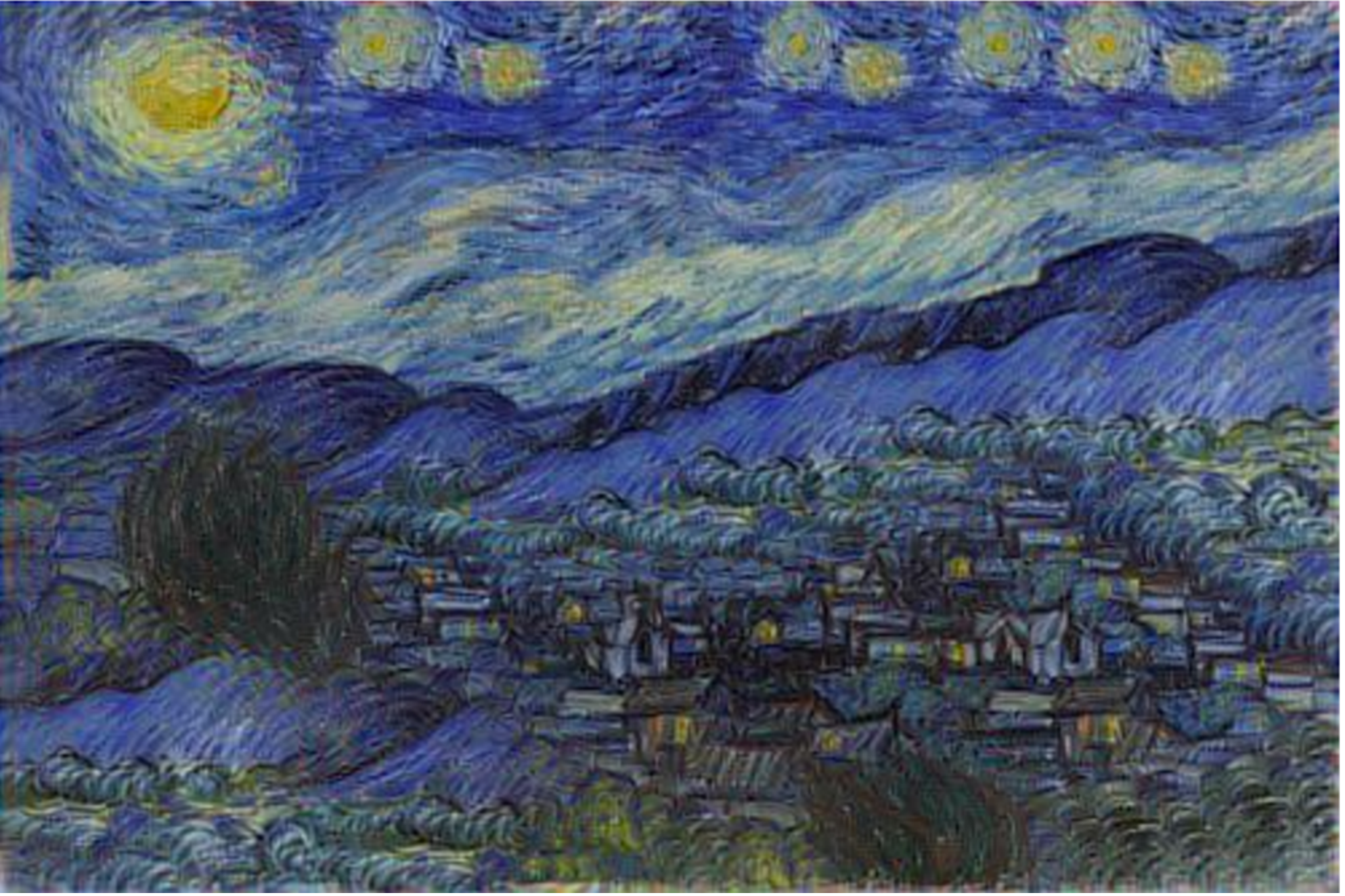
Image Analogies

Aaron Hertzmann
Charles Jacobs
Nuria Oliver
Brian Curless
David Salesin

Inter-modal: Style Transfer ("Style Net" 2015)

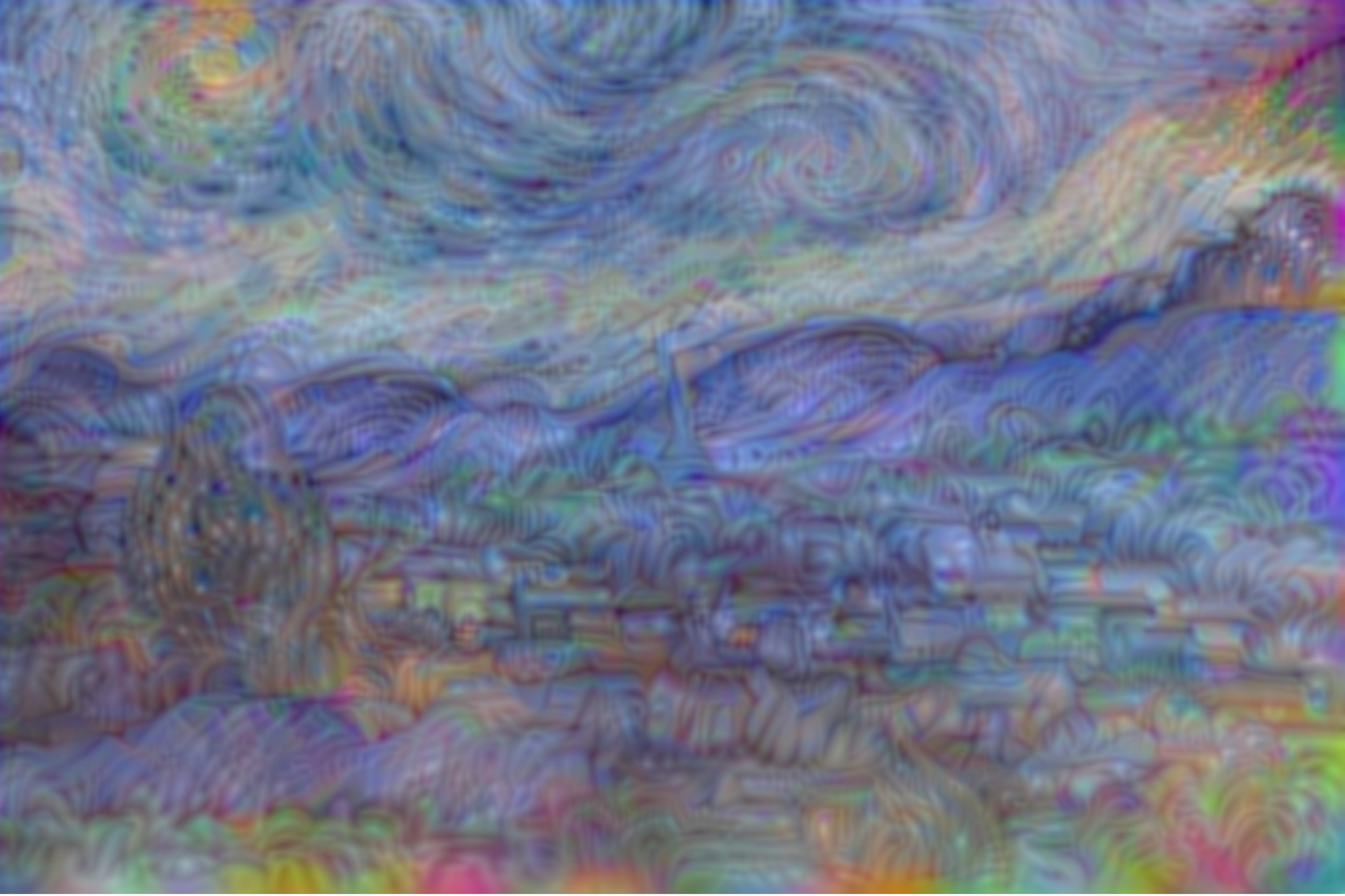


Leon A. Gatys, Alexander S. Ecker, Matthias Bethge , 2015.
A Neural Algorithm of Artistic Style (GitXiv)

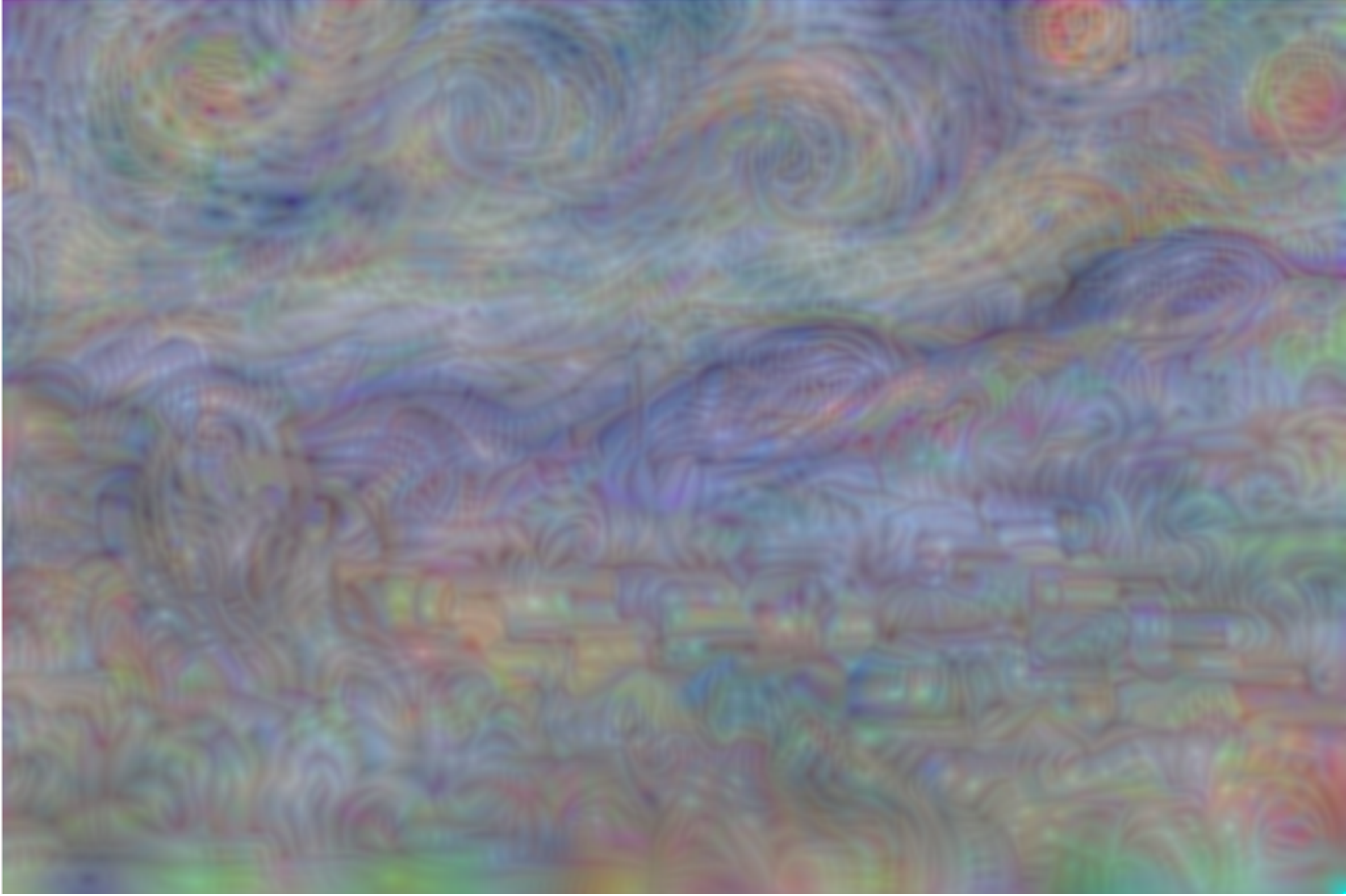


style layers: 3_1,4_1,5_1

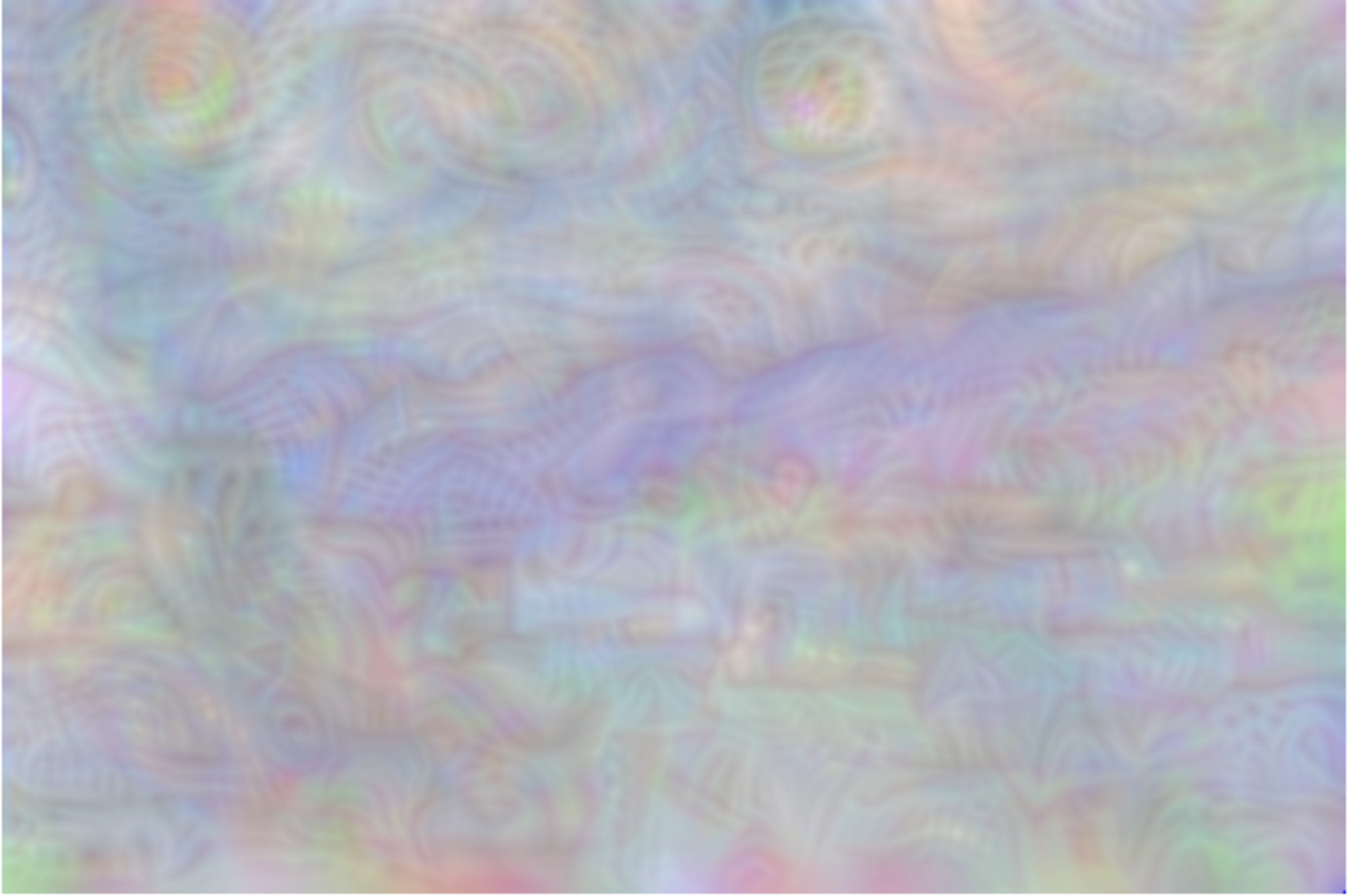
style layers: 3_2



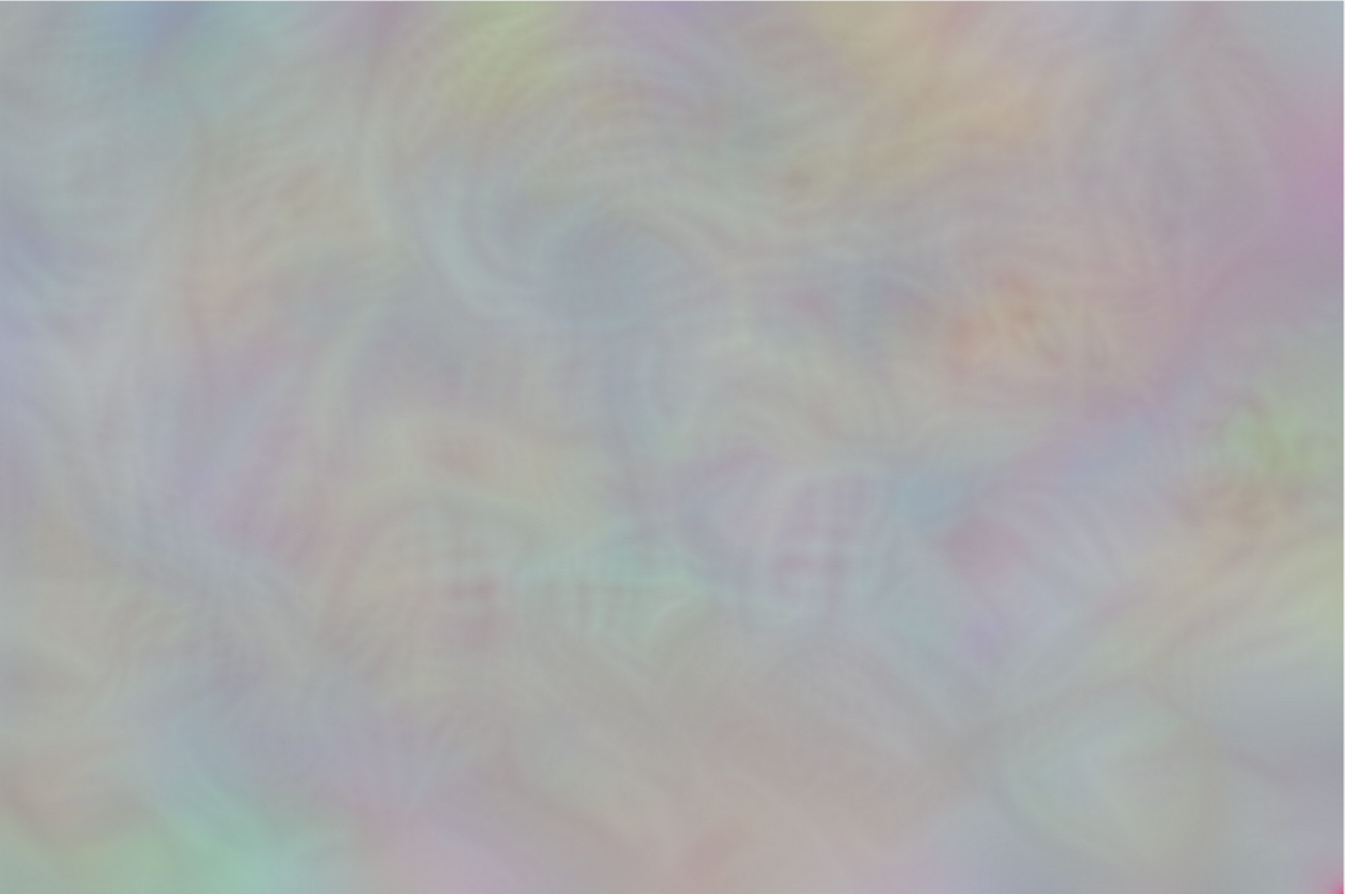
style layers: 5_1



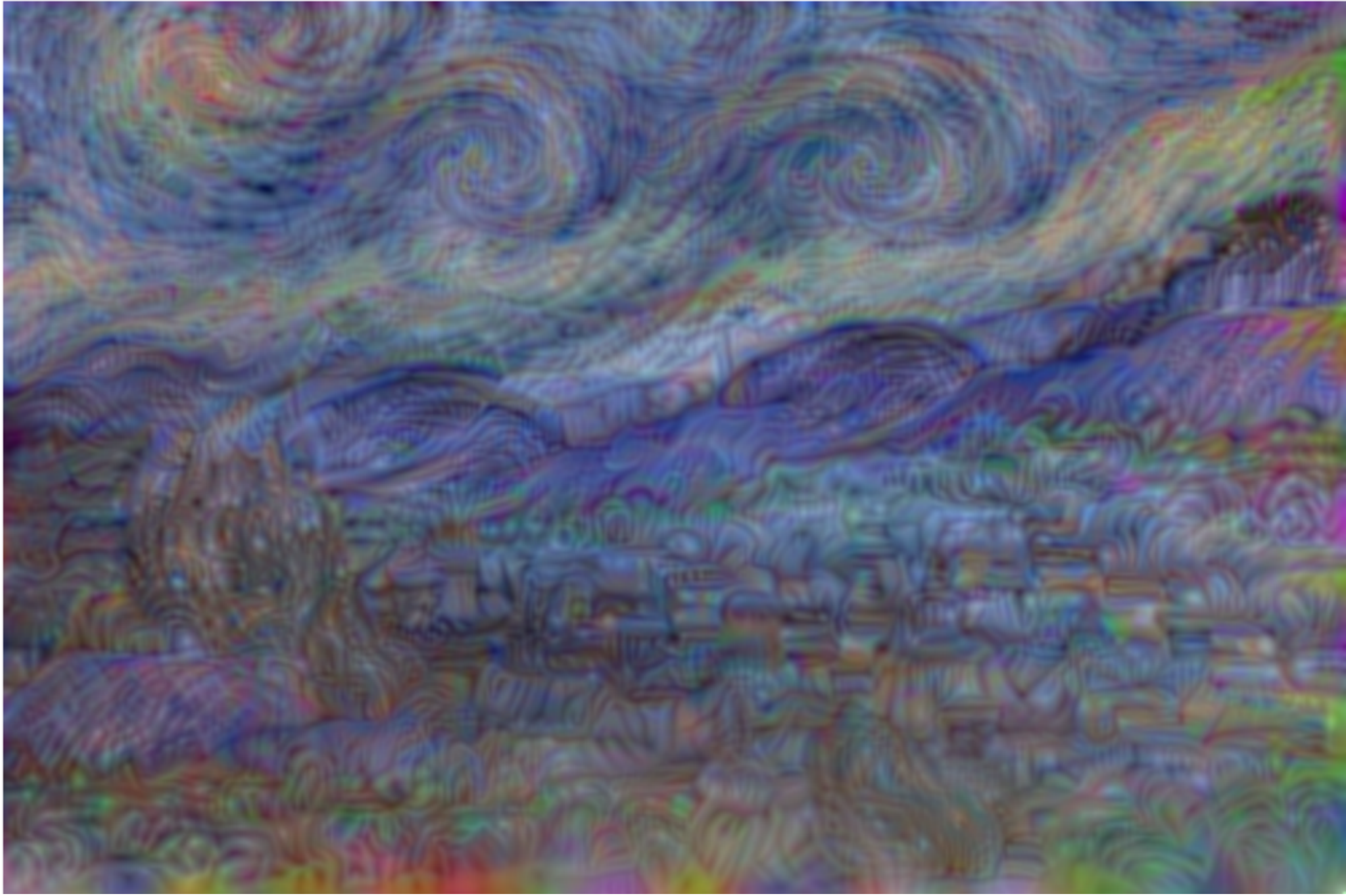
style layers: 5_2



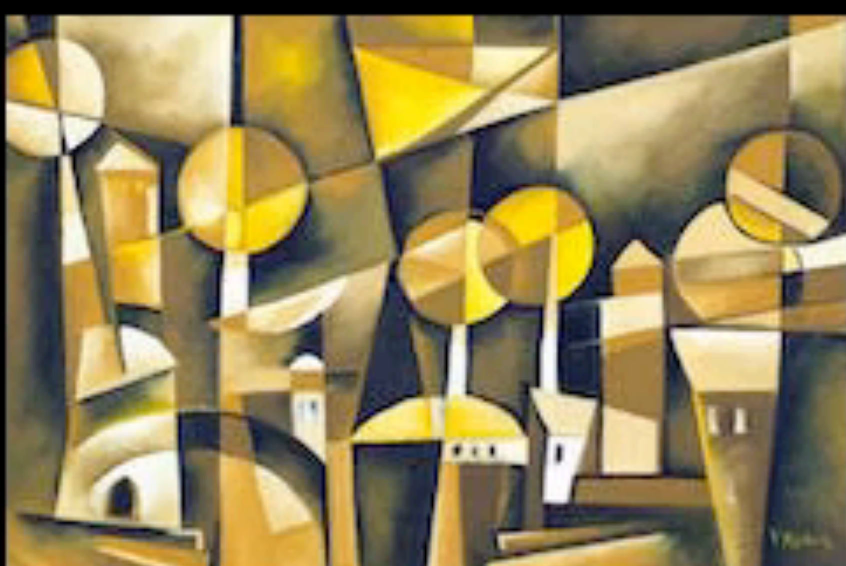
style layers: 5_3



style layers: 5_4



style layers: 5_1 + 5_2 + 5_3 + 5_4



Pablo Picasso

Gene Kogan, 2015. Why is a Raven Like a Writing Desk? (vimeo)

Inter-modal: Style Transfer+MRF (2016)

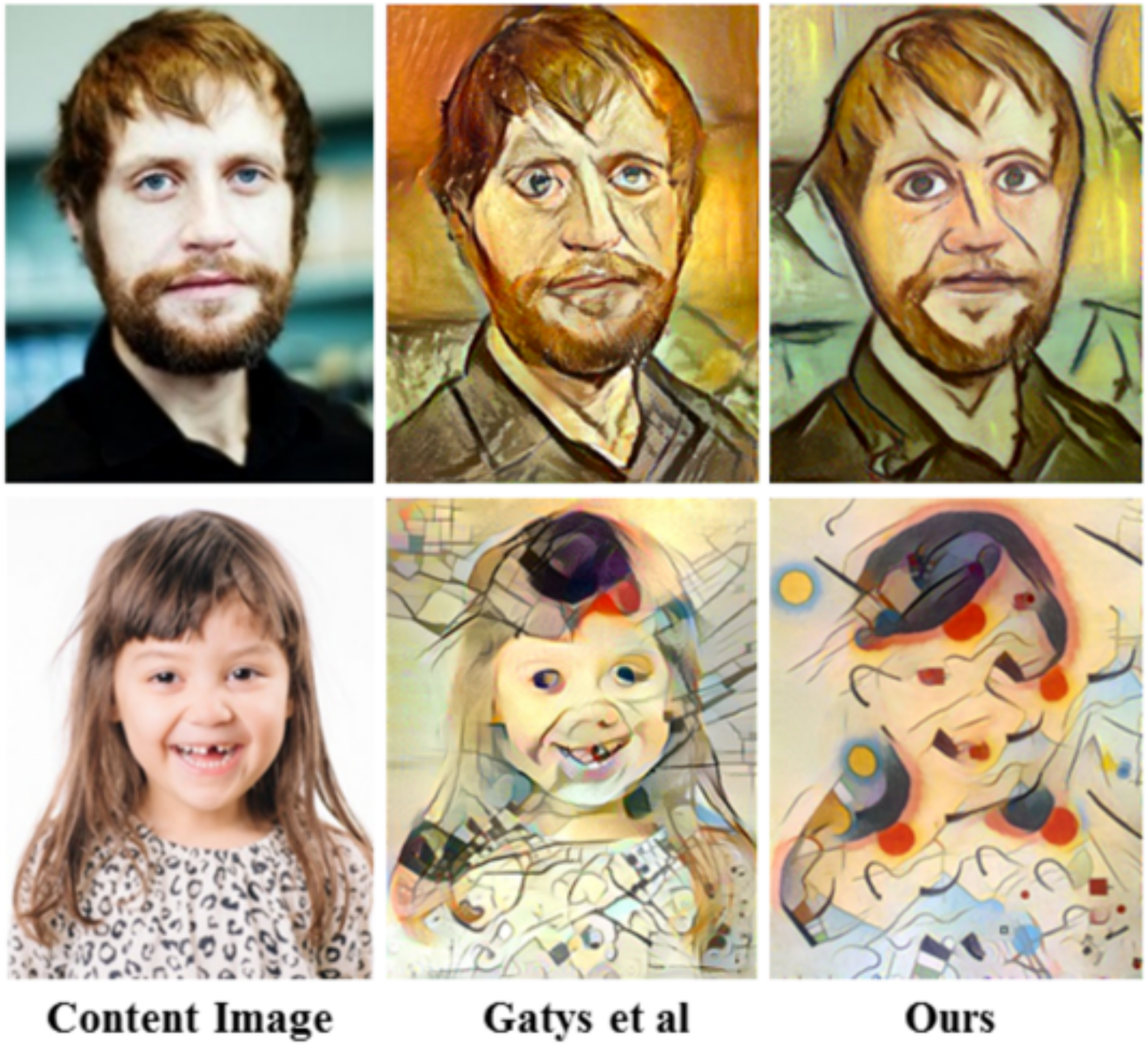


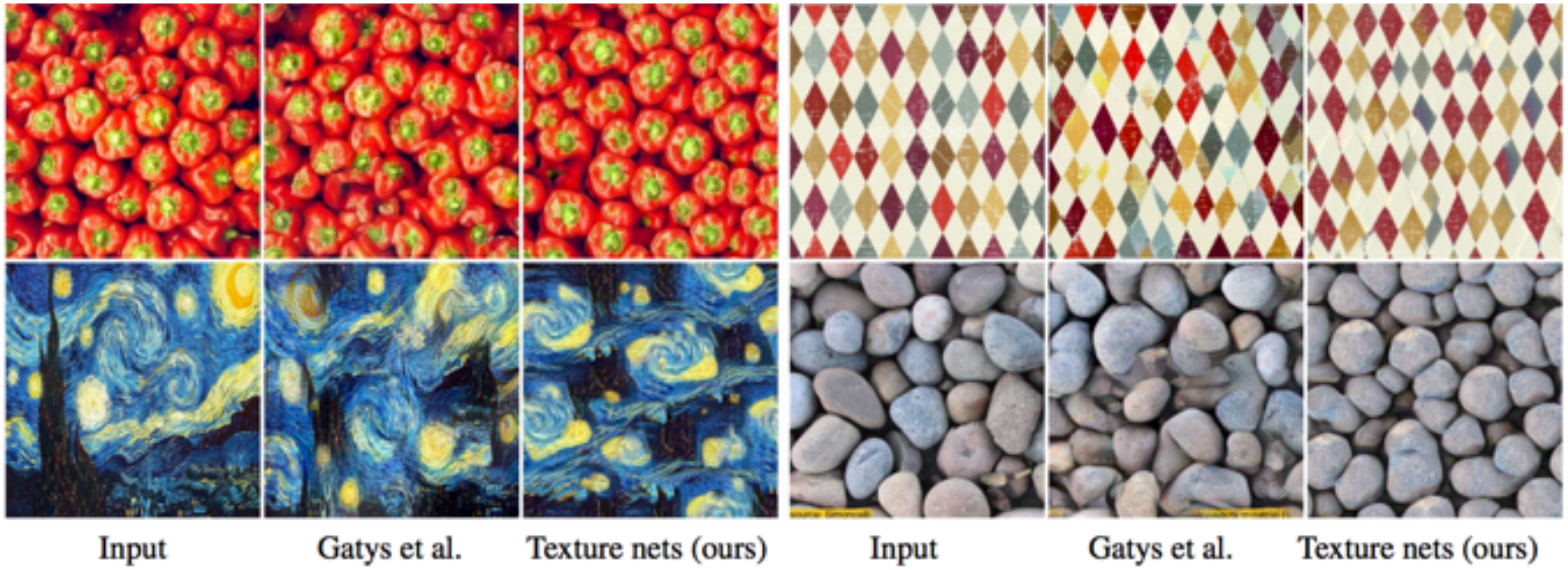
Figure 6: Comparison with Gatys et al. [8] for artistic synthesis. Content images credited to flickr users *Christopher Michel* (top) and *Peter Dahlgren* (bottom).

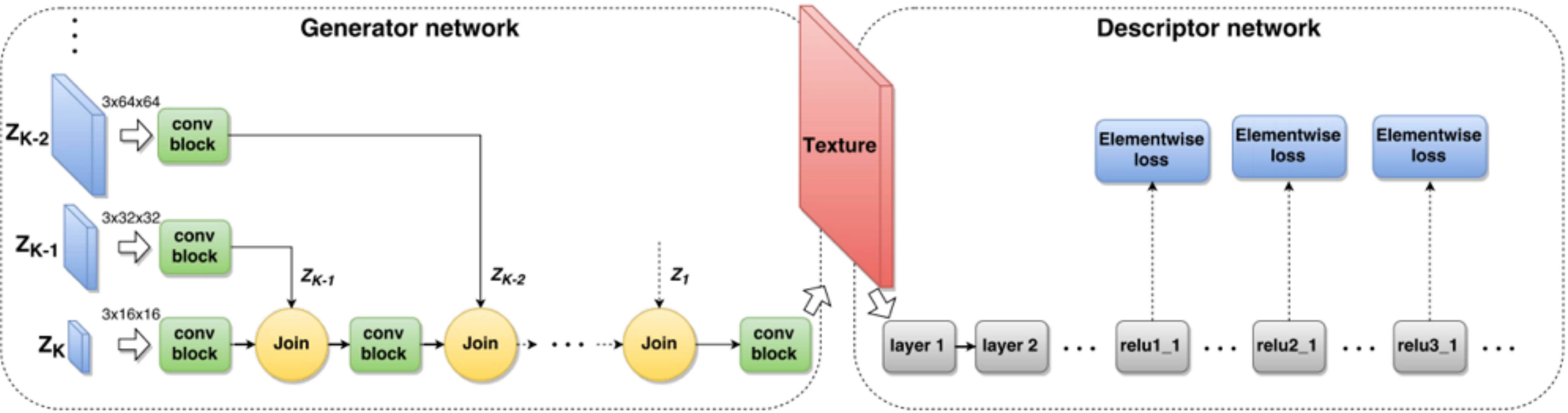
Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis, 2016, Chuan Li, Michael Wand

Inter-modal: Style Transfer+MRF (2016)

Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis, 2016, Chuan Li, Michael Wand

Inter-modal: Pretrained Style Transfer (2016)

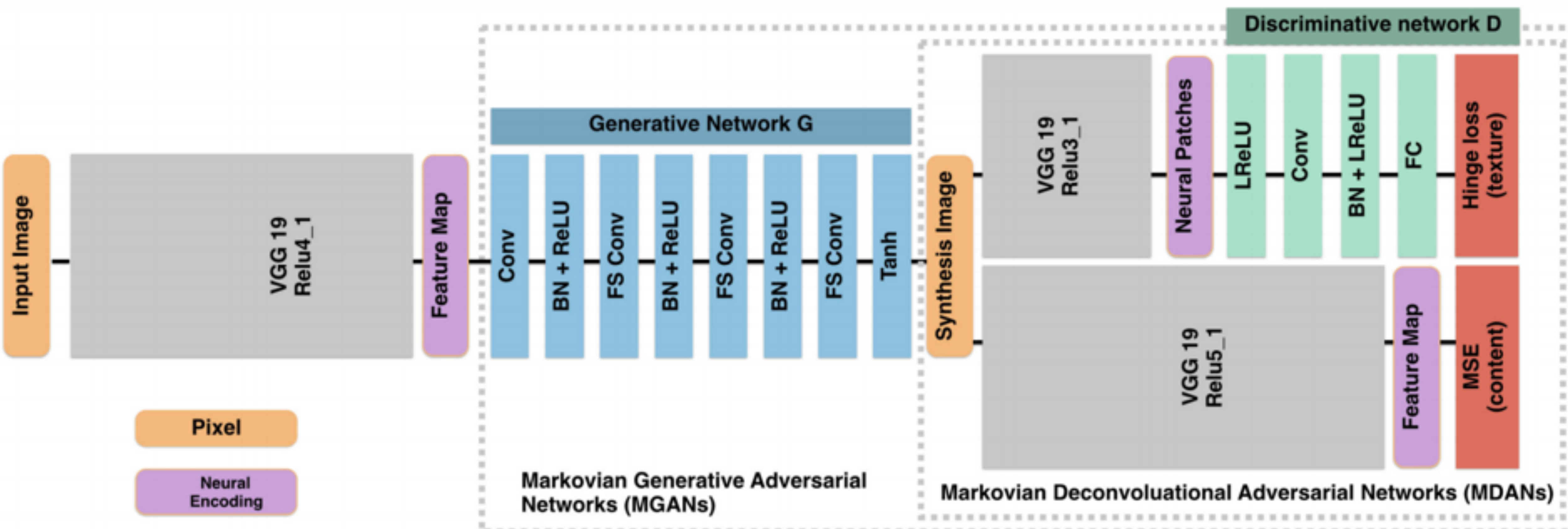




500x speedup! (avg min loss from 10s to 20ms)

Texture Networks: Feed-forward Synthesis of Textures and Stylized Images, 2016, Dmitry Ulyanov, Vadim Lebedev, Andrea Vedaldi, Victor Lempitsky

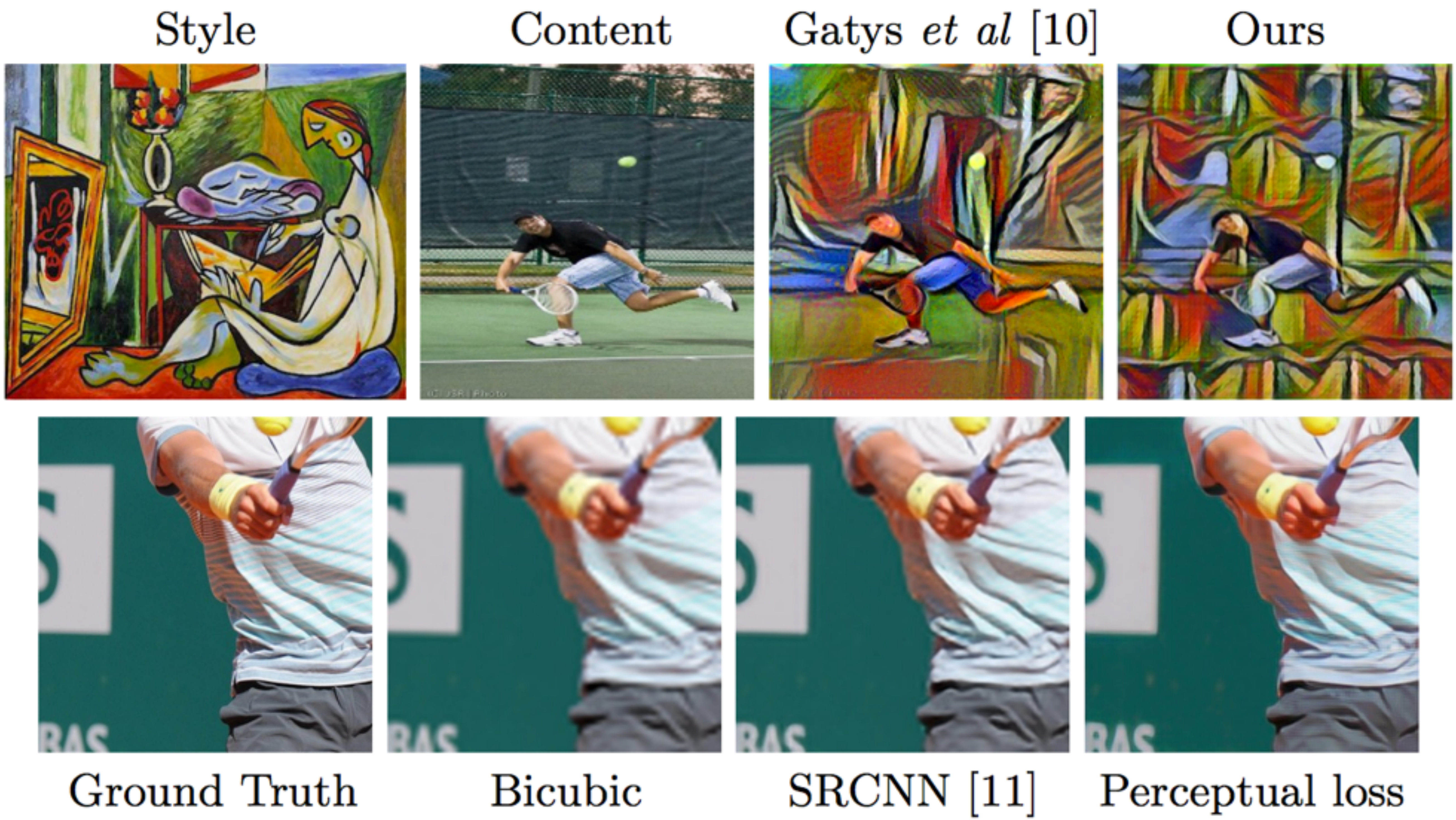
Inter-modal: Pretrained Style Transfer #2 (2016)



similarly 500x speedup

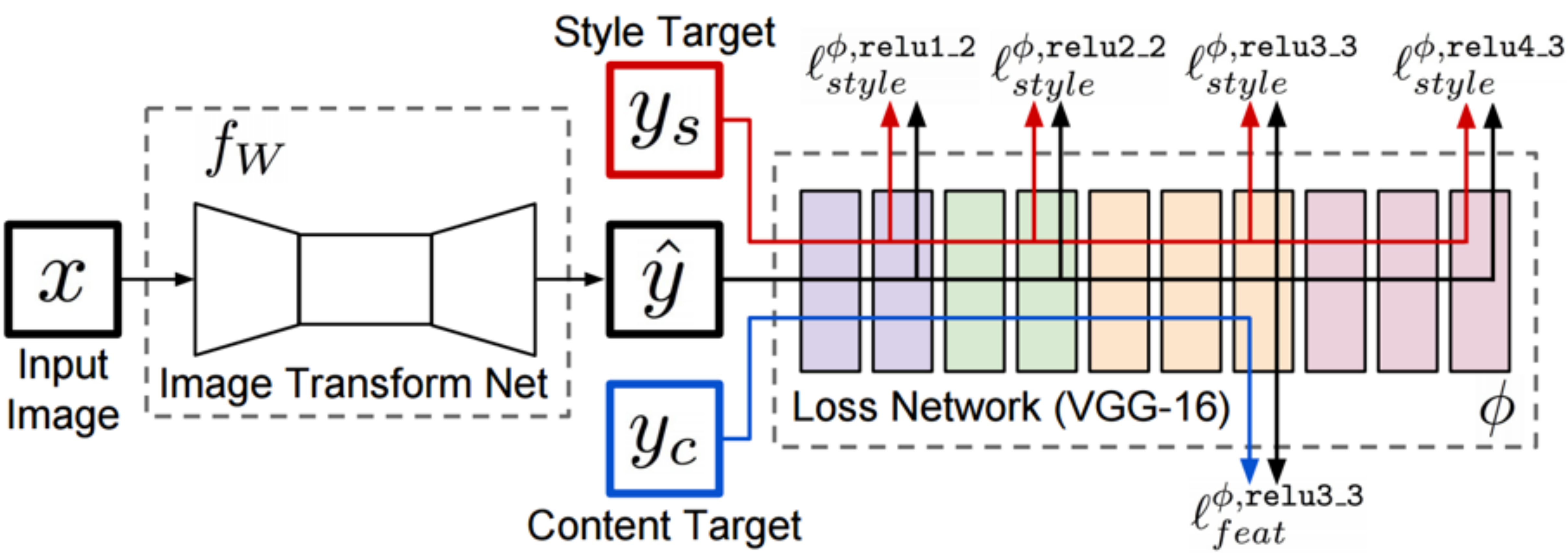
Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, Chuan Li, Michael Wand, 2016

Inter-modal: Perceptual Loss ST (2016)



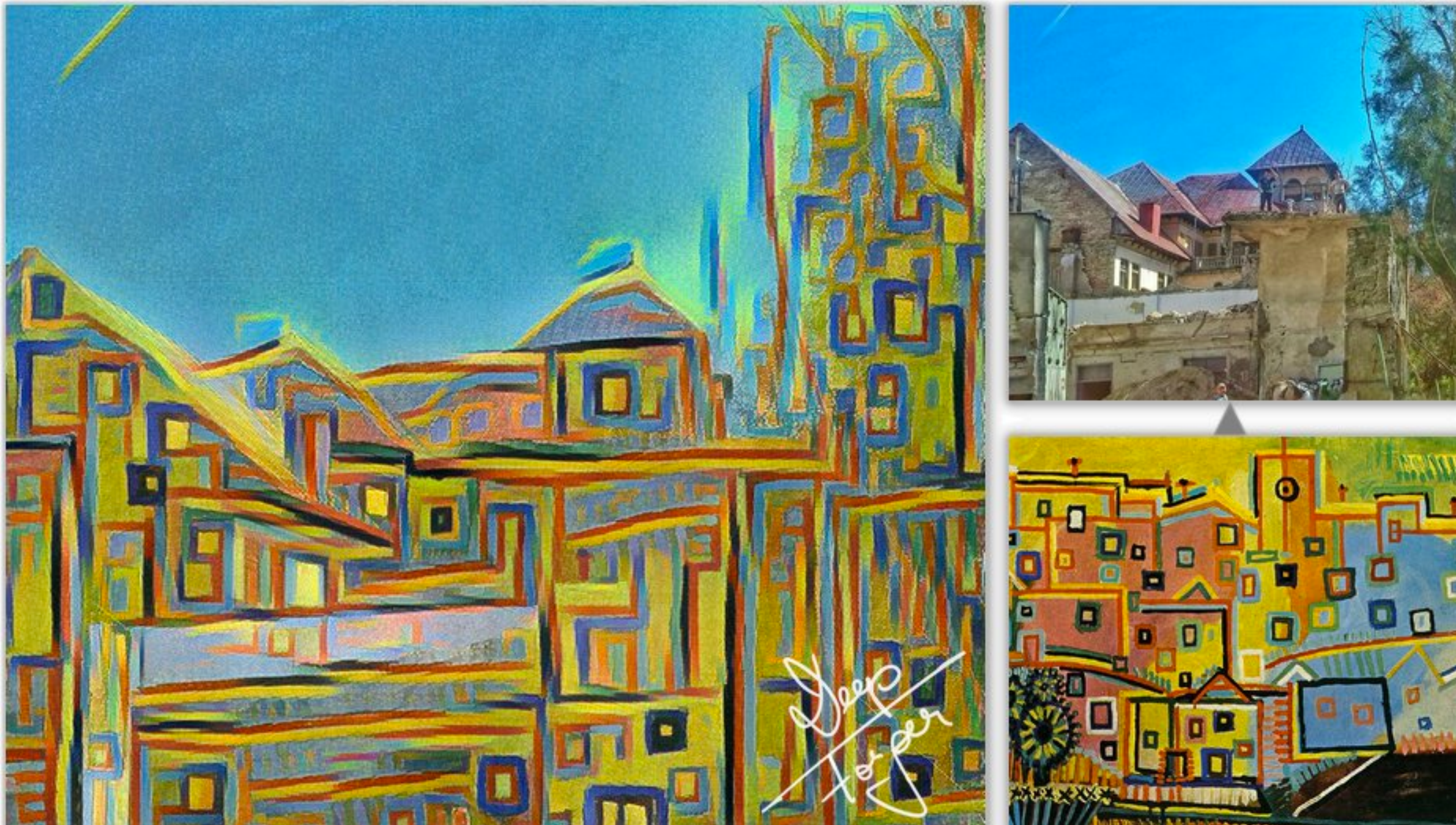
Perceptual Losses for Real-Time Style Transfer and Super-Resolution, 2016, Justin Johnson, Alexandre Alahi, Li Fei-Fei

	Ground Truth	Bicubic	Ours (ℓ_{pixel})	SRCNN [11]	Ours (ℓ_{feat})
This image		31.78 / 0.8577	31.47 / 0.8573	32.99 / 0.8784	29.24 / 0.7841
Set5 mean		28.43 / 0.8114	28.40 / 0.8205	30.48 / 0.8628	27.09 / 0.7680



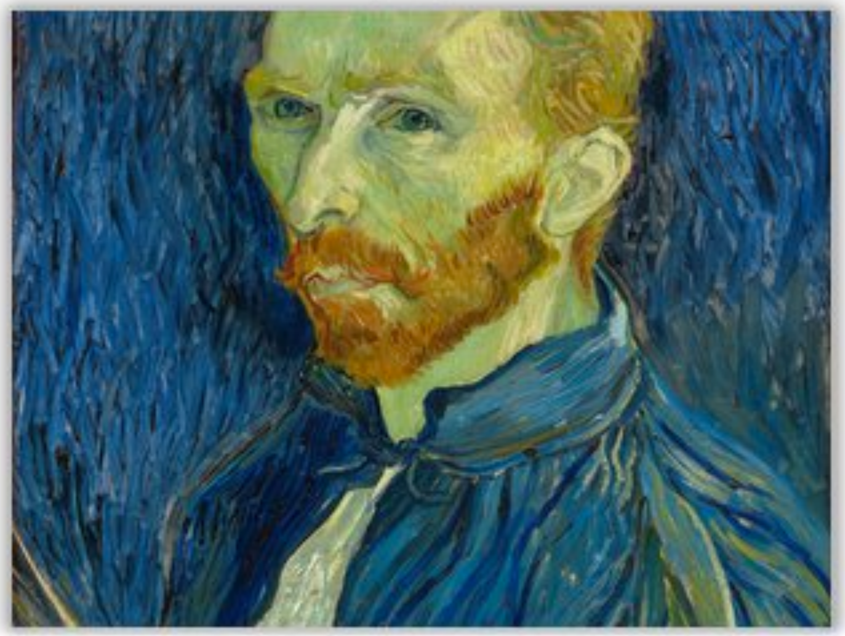
Perceptual Losses for Real-Time Style Transfer and Super-Resolution, 2016, Justin Johnson, Alexandre Alahi, Li Fei-Fei

Inter-modal: Style Transfer ("Style Net" 2015)



@DeepForger

Inter-modal: Style Transfer ("Style Net" 2015)



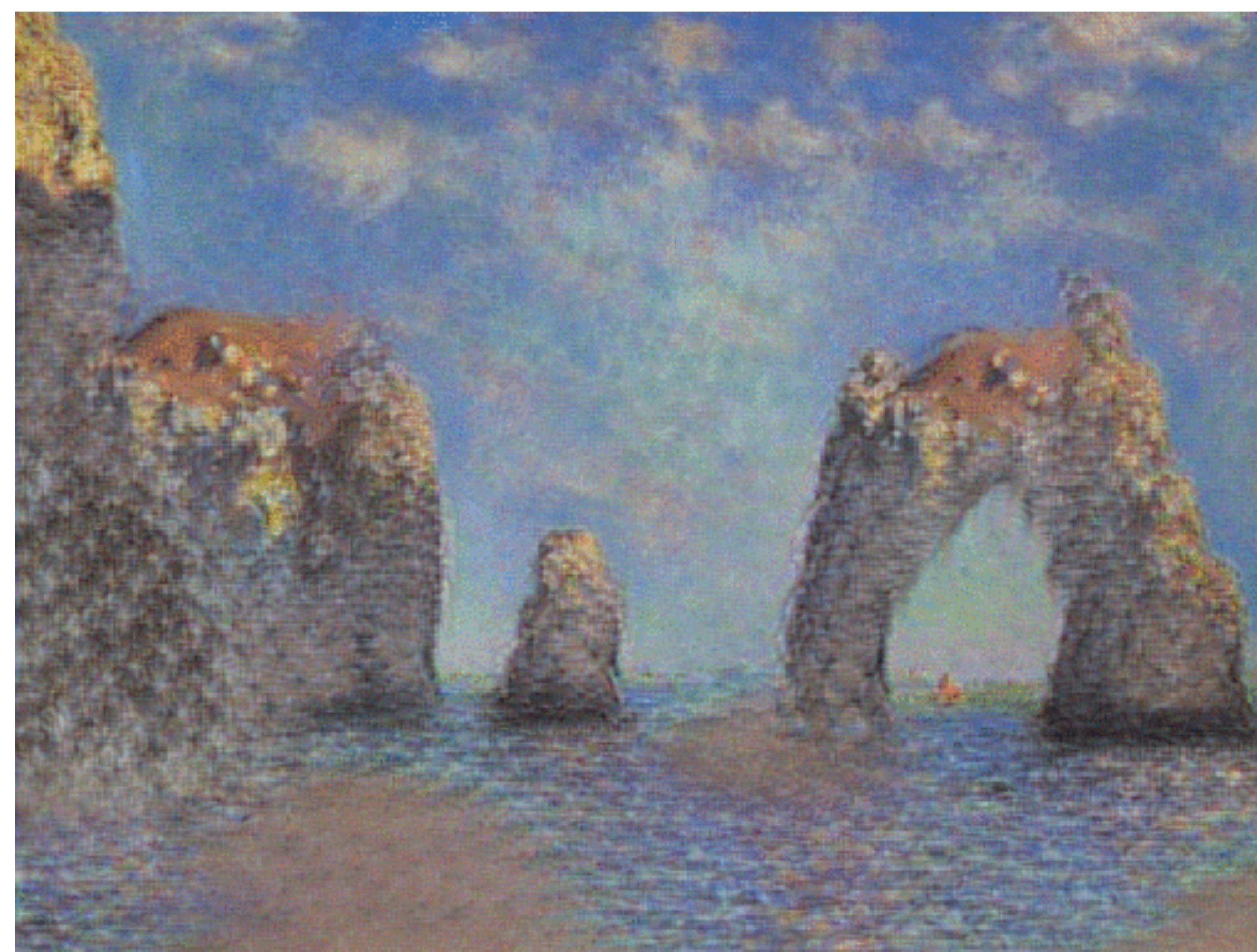
@DeepForger

Inter-modal: Semantic Style Transfer

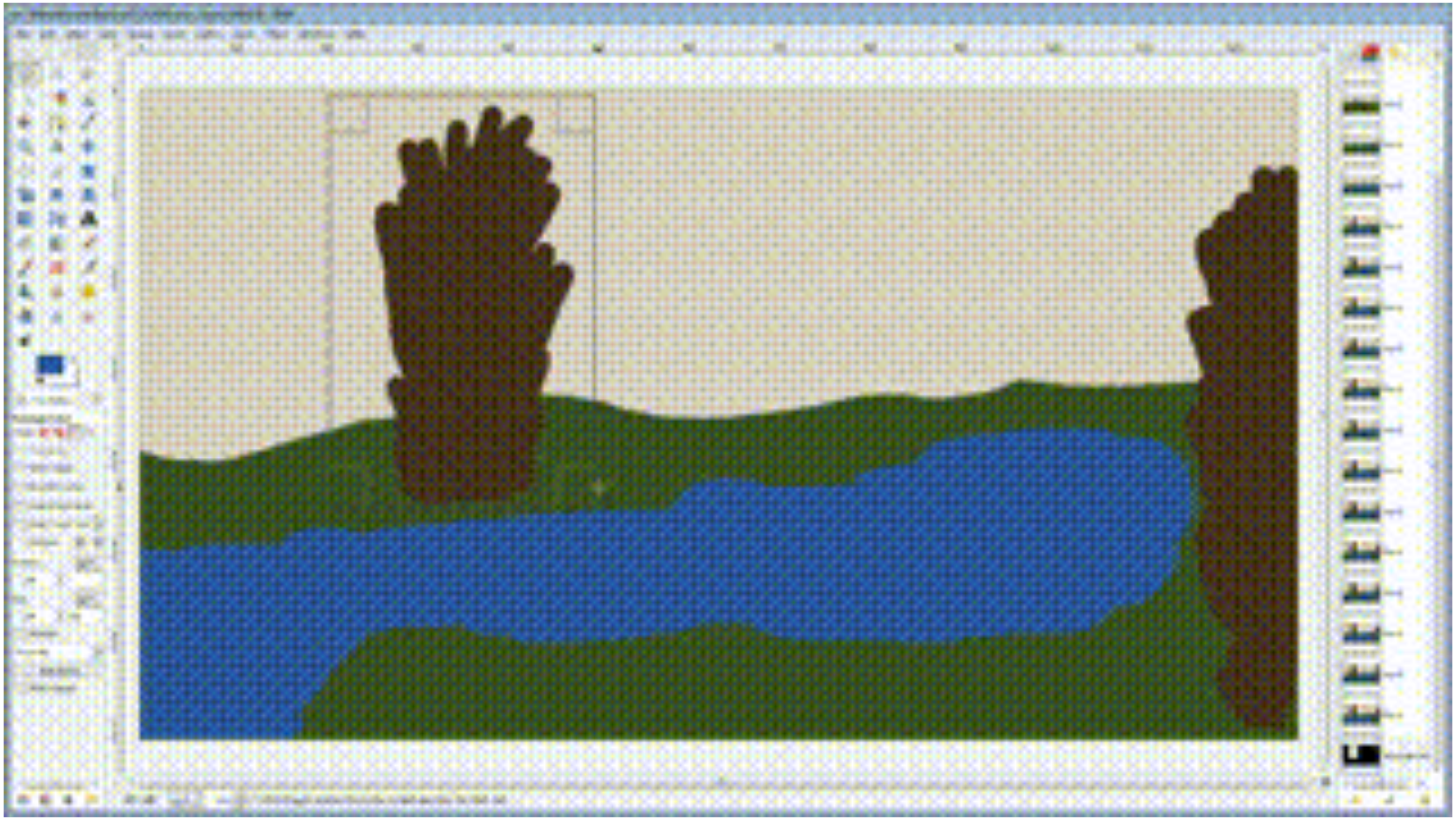
+

+

=



Semantic Style Transfer (“Neural Doodle”)



https://github.com/alexjc/neural-doodle

Synthesize textures

Original Texture

Synthesized Texture #1

Synthesized Texture #2

Synthesise textures (random weights)

Random Weights (kinda like "Extreme Learning Machines")

- which activation function should i use?
- pooling?
- min nr of units?

Experiment:

```
48 units,  3x3 shape                  # conv1_1
48 units,  3x3 shape                  # conv1_2
80 units,  2x2 shape, stride 2x2      # conv2_1
80 units,  3x3 shape                  # conv2_2
112 units, 2x2 shape, stride 2x2      # conv3_1
112 units, 3x3 shape                  # conv3_2
112 units, 3x3 shape                  # conv3_3
176 units, 2x2 shape, stride 2x2      # conv4_1
176 units, 3x3 shape                  # conv4_2
176 units, 3x3 shape                  # conv4_3
```

https://nucl.ai/blog/extreme-style-machines/

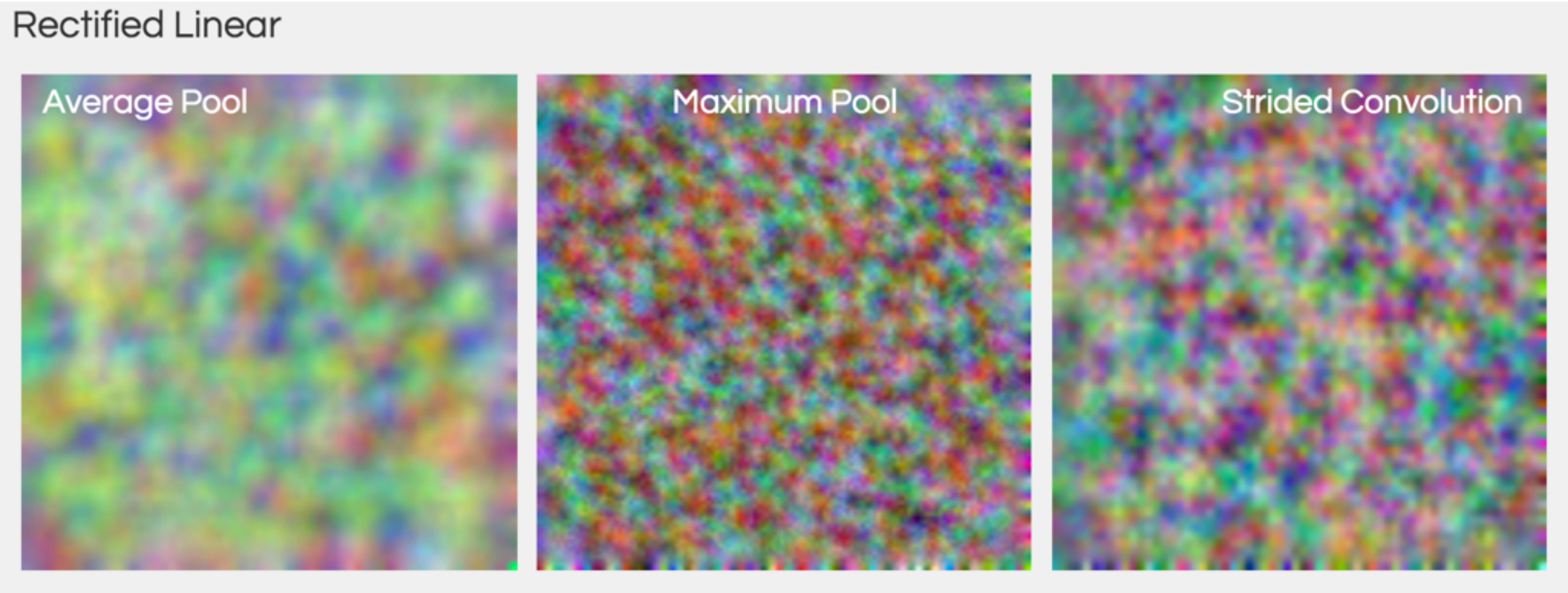
Synthesise textures (random weights)

totally random initialised weights:

https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

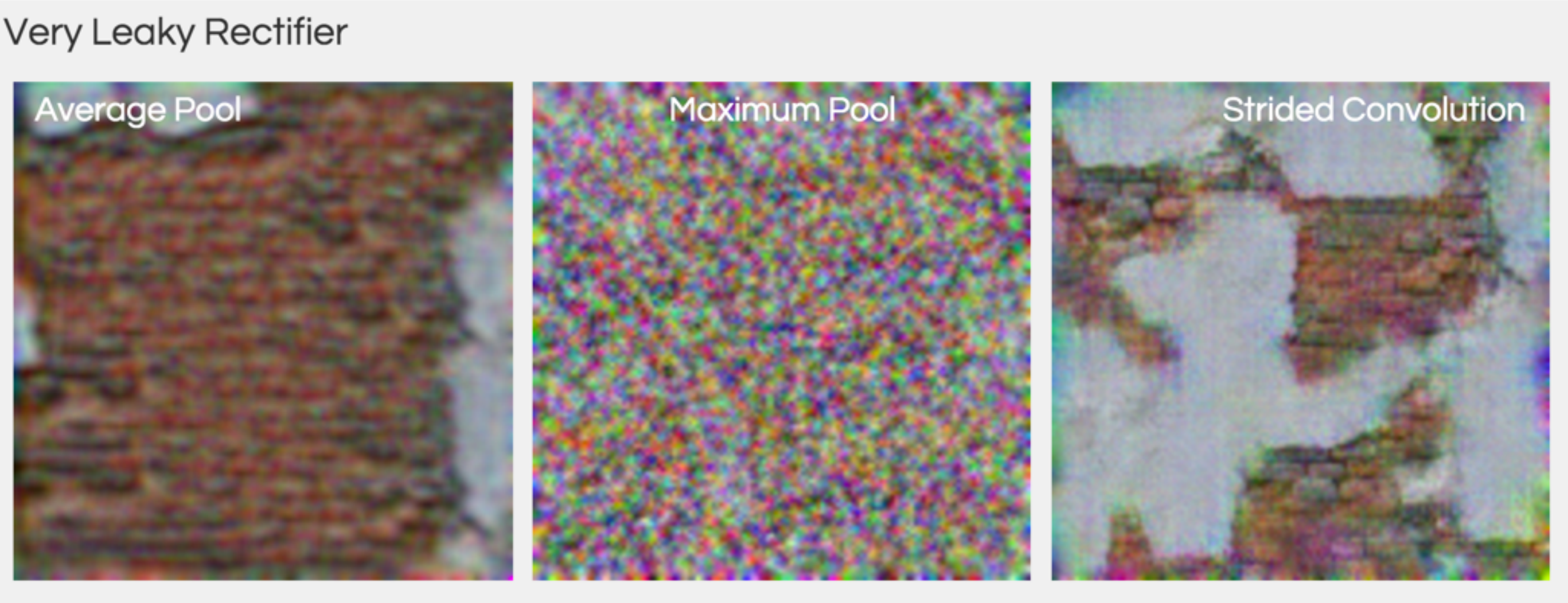
Activation Functions



https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

Activation Functions



https://nucl.ai/blog/extreme-style-machines/

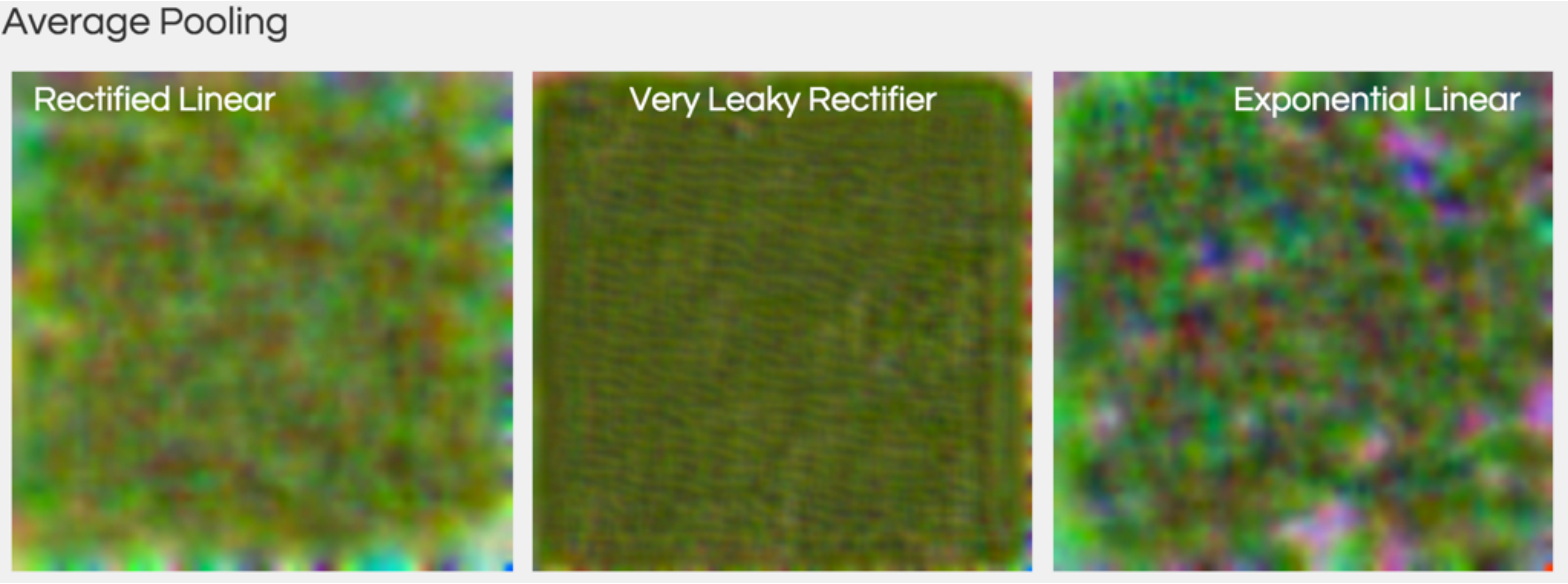
Synthesise textures (randim weights)

Activation Functions

https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

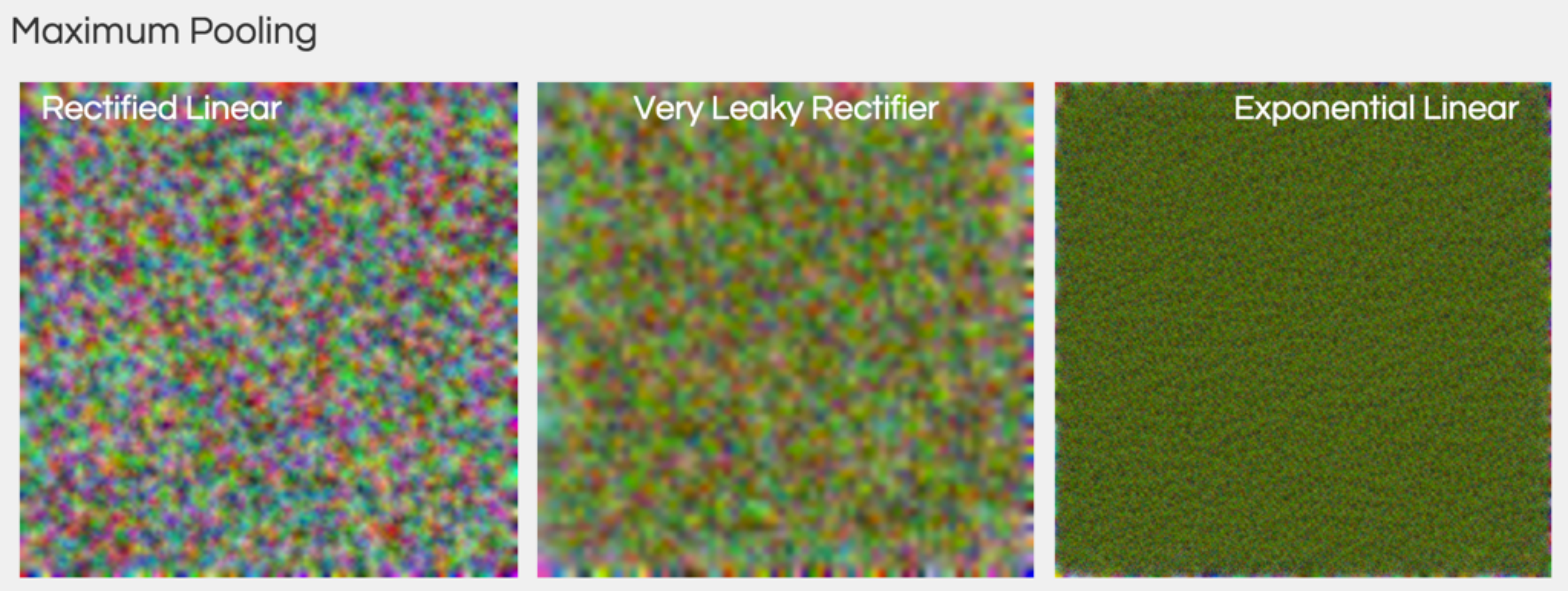
Down-sampling



https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

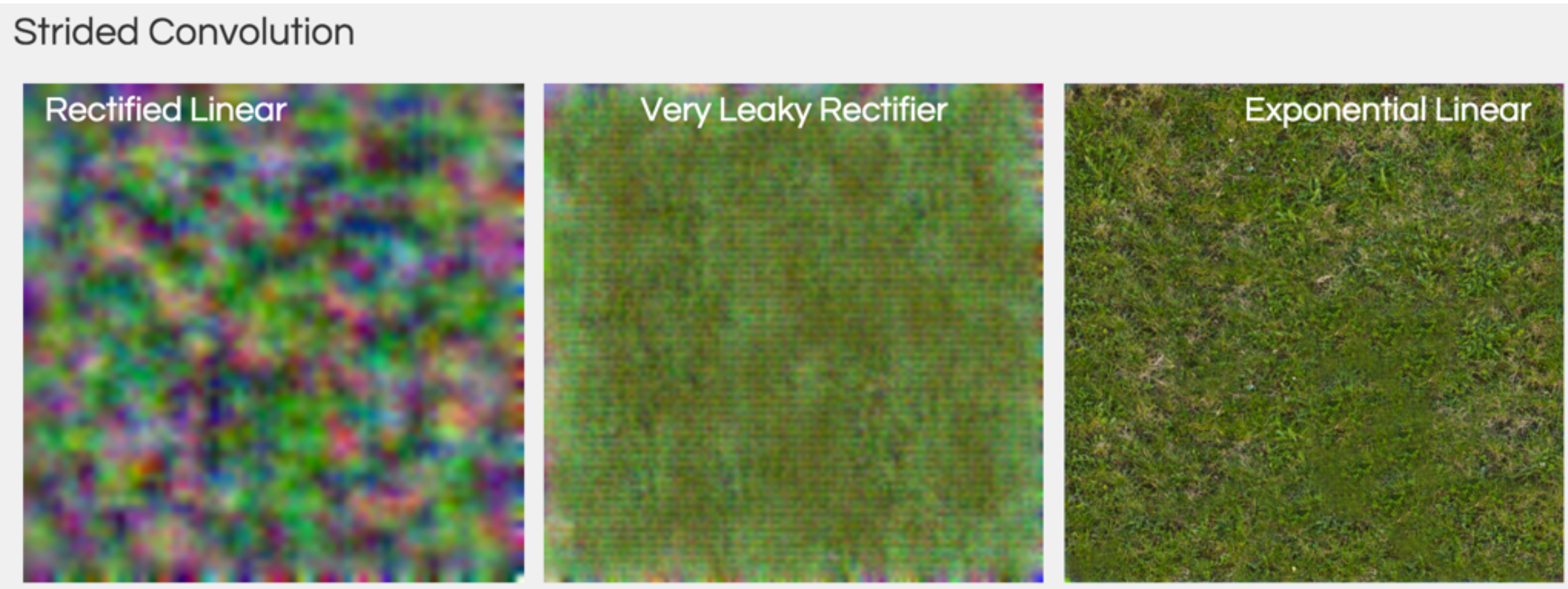
Down-sampling



https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

Down-sampling



https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

Nr of units

https://nucl.ai/blog/extreme-style-machines/

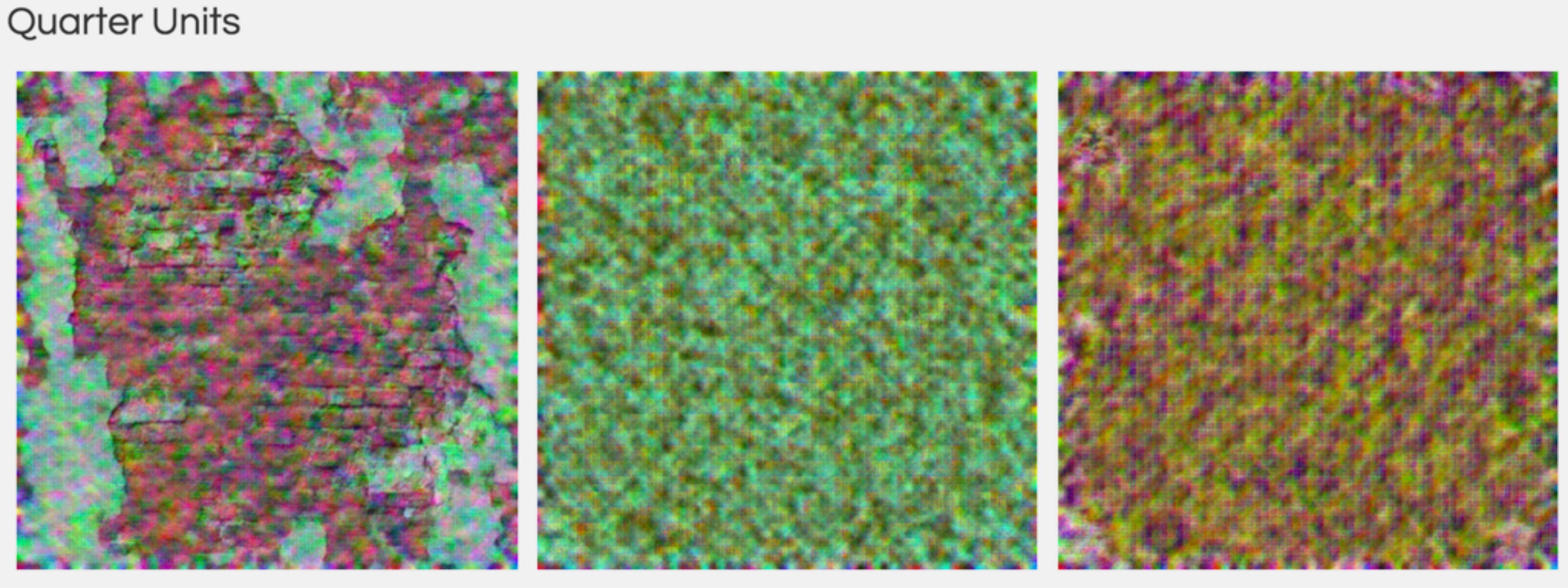
Synthesise textures (randim weights)

Nr of units

https://nucl.ai/blog/extreme-style-machines/

Synthesise textures (randim weights)

Nr of units



https://nucl.ai/blog/extreme-style-machines/

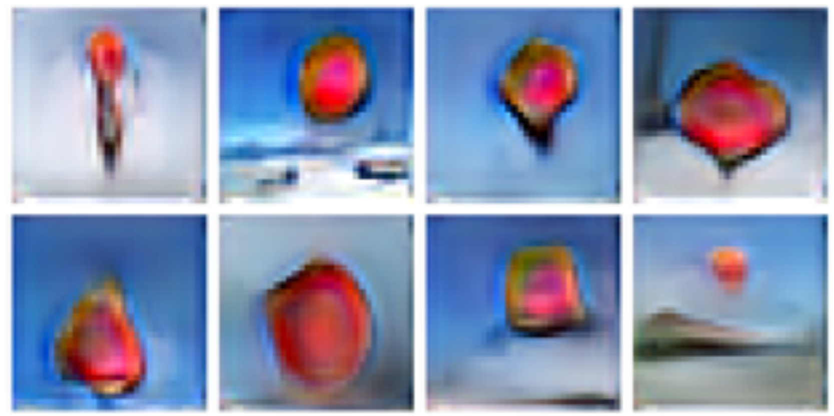
"Style Transfer" papers

- Image Analogies, **<u>2001</u>**, A. Hertzmann, C. Jacobs, N. Oliver, B. Curless, D. Sales
- A Neural Algorithm of Artistic Style, 2015. Leon A. Gatys, Alexander S. Ecker, Matthias Bethge
- Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis, 2016, Chuan Li, Michael Wand
- Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artworks, 2016, Alex J. Champandard
- Texture Networks: Feed-forward Synthesis of Textures and Stylized Images, 2016, Dmitry Ulyanov, Vadim Lebedev, Andrea Vedaldi, Victor Lempitsky
- Perceptual Losses for Real-Time Style Transfer and Super-Resolution, 2016, Justin Johnson, Alexandre Alahi, Li Fei-Fei
- Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, 2016, Chuan Li, Michael Wand

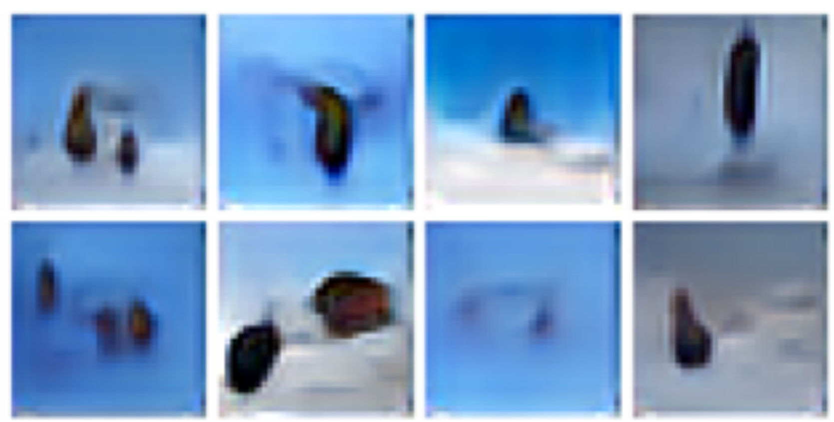
@DeepForger

Caption -> Image generation

"A stop sign is flying in blue skies."



"A herd of elephants flying in the blue skies."



Elman Mansimov, Emilio Parisotto, Jimmy Lei Ba, Ruslan Salakhutdinov, 2015. Generating Images from Captions with Attention (arxiv) (examples)

Image Colorisation

More...

more at:
http://gitxiv.com/category/computer-vision
http://www.creativeai.net/ (no category for images just yet)

Audio Generation

Early LSTM music composition (2002)

Douglas Eck and Jurgen Schmidhuber (2002) Learning The Long-Term Structure of the Blues?

Markov constraints

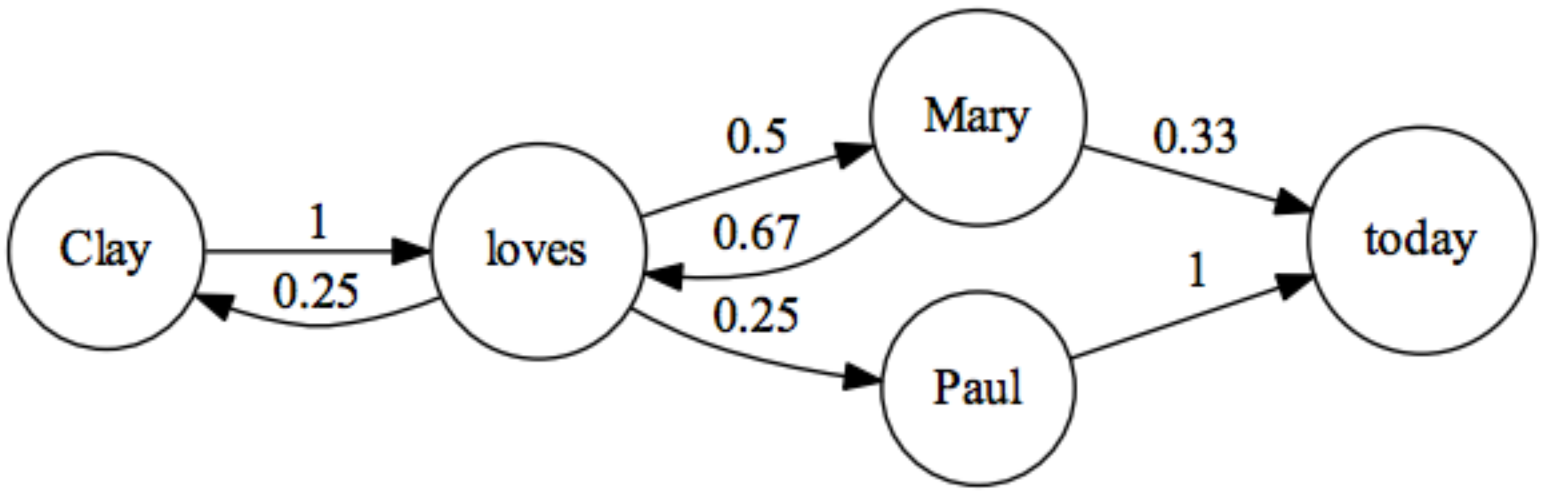


Figure 1. An order-1 Markov process learned from a corpus composed of five words.

$$P\left(s_i | s_1, \cdots, s_{i-1}\right) = P\left(s_i | s_{i-1}\right)$$

http://www.flow-machines.com/

Ode to Joy in several styles

http://www.flow-machines.com/

Audio Generation: Midi

+ "modded VRNN:

A Recurrent Latent Variable Model for Sequential Data, 2016,
J. Chung, K. Kastner, L. Dinh, K. Goel, A. Courville, Y. Bengio

https://soundcloud.com/graphific/pyotr-lstm-tchaikovsky

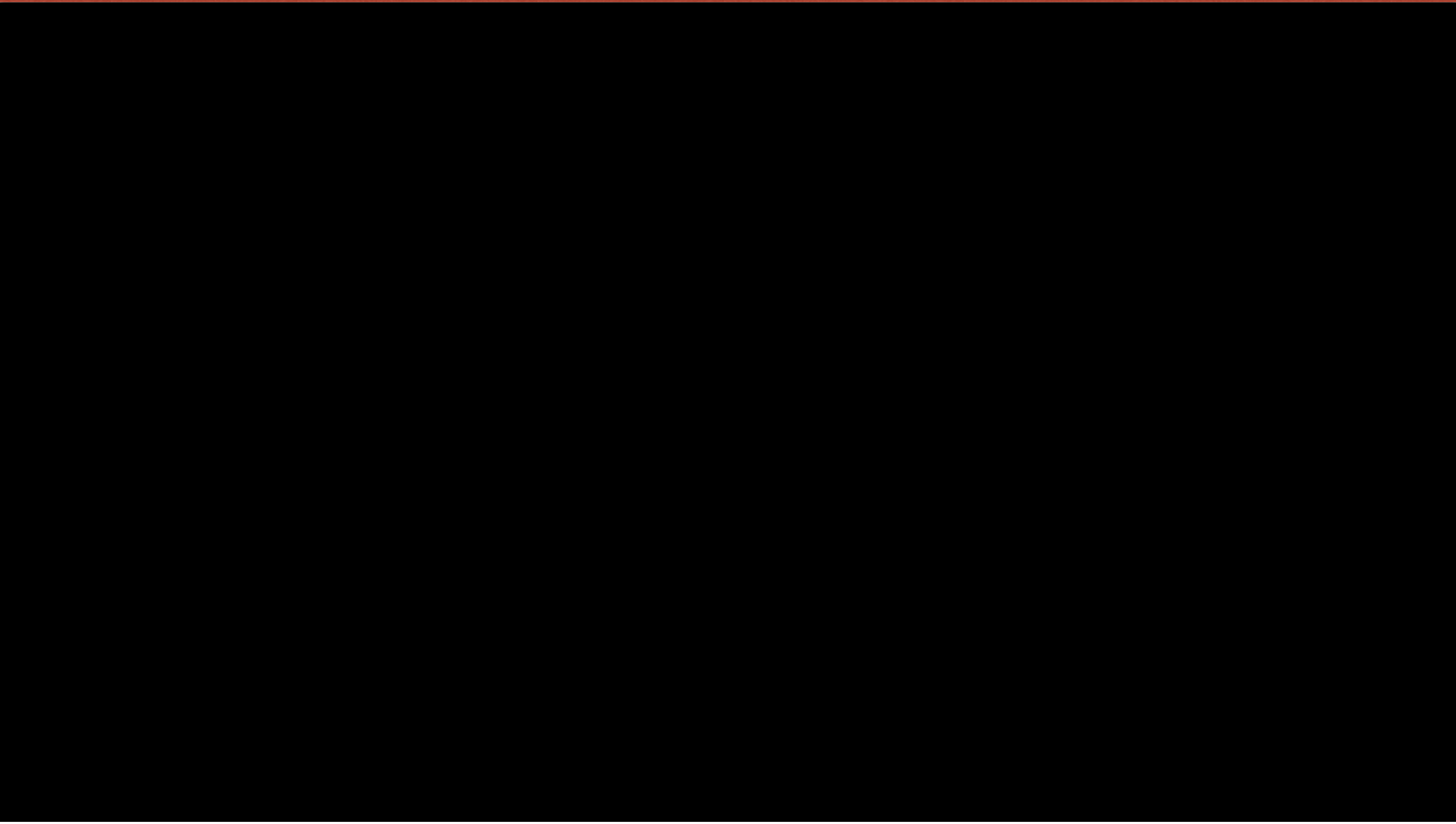
Audio Generation: Midi

+ "modded VRNN:

A Recurrent Latent Variable Model for Sequential Data, 2016,
J. Chung, K. Kastner, L. Dinh, K. Goel, A. Courville, Y. Bengio

https://soundcloud.com/graphific/neural-remix-net

Audio Generation: Raw



stanford cs224d project   Gated Recurrent Unit (GRU)

Aran Nayebi, Matt Vitelli (2015) GRUV: Algorithmic Music Generation using Recurrent Neural Networks

LSTM improvements

- LSTM improvements
  - **Recurrent Batch Normalization** http://gitxiv.com/posts/MwSDm6A4wPG7TcuPZ/recurrent-batch-normalization
    - also: hidden-to-hidden transition (earlier only input-to-hidden transformation of RNNs)
    - faster convergence and improved generalization.

LSTM improvements

- LSTM improvements

  - **weight normalisation** http://gitxiv.com/posts/p9B6i9Kzbkc5rP3cp/weight-normalization-a-simple-reparameterization-to

LSTM improvements

- LSTM improvements

  - **Associative Long Short-Term Memory** http://gitxiv.com/posts/jpfdiFPsu5c6LLsF4/associative-long-short-term-memory

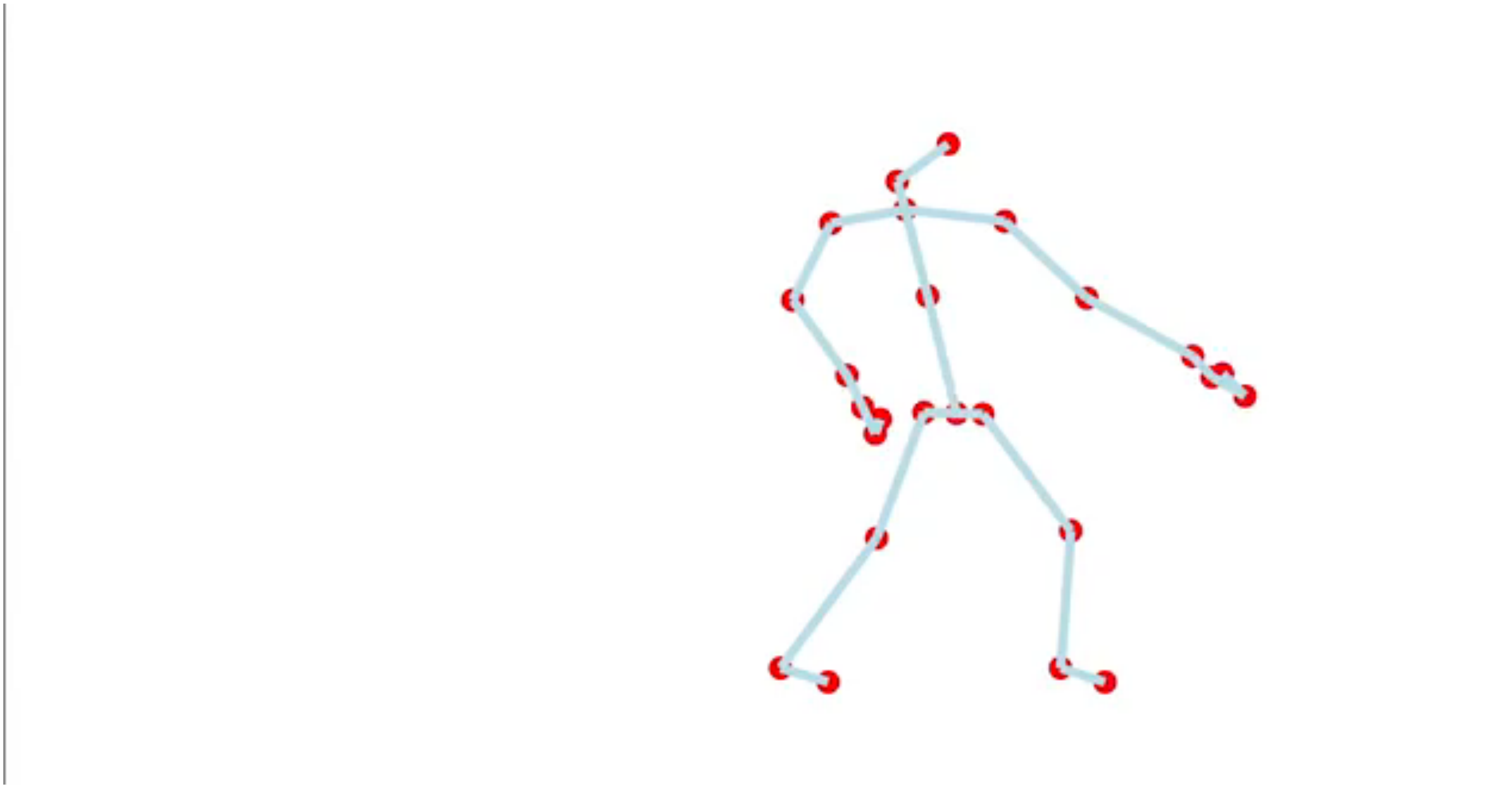
LSTM improvements

- LSTM improvements
  - **Bayesian RNN dropout** http://gitxiv.com/posts/CsCDjy7WpfcBvZ88R/bayesianrnn

Continuous Generation

Chor-RNN

$\cancel{t_t = x_{t+1}}$



Luka Crnkovic-Friis & Louise Crnkovic-Friis (2016) Generative Choreography using Deep Learning

Generative

- Mixture Density LSTM

1 $$p(\boldsymbol{t}|\boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i(\boldsymbol{x})\varphi_i(\boldsymbol{t}|\boldsymbol{x})$$

2 $$\varphi_i(\boldsymbol{t}|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{c}{2}}\sigma_i(\boldsymbol{x})^c} e^{-\frac{\|\boldsymbol{t}-\boldsymbol{\mu}_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}}$$

3 $$\boldsymbol{z} = [z_1^{\alpha}, \ldots, z_m^{\alpha}, z_{m+1}^{\mu}, \ldots, z_{mc+m+1}^{\mu}, z_{mc+m+2}^{\sigma}, \ldots, z_{m(c+2)}^{\sigma}]$$

4 $$E^q = -\log\left[\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\varphi_i(\boldsymbol{t}^q|\boldsymbol{x}^q)\right]$$

5 $$\alpha_i = \frac{e^{z_i^{\alpha}}}{\sum_{j=1}^{M} e^{z_j^{\alpha}}}, \quad \sigma_i = e^{z_i^{\sigma}}, \quad \mu_i = z_{ik}^{\mu}$$

$$\frac{\partial E^q}{\partial z_k^{\alpha}} = \alpha_k - \frac{\alpha_k \varphi_k}{\sum_{j=1}^{m} \alpha_j \varphi_j}$$

$$\frac{\partial E^q}{\partial z_{ik}^{\mu}} = \frac{\alpha_i \varphi_i}{\sum_{j=1}^{m} \alpha_j \varphi_j} \frac{\mu_{ik} - t_k}{\sigma_i^2}$$ 6

$$\frac{\partial E^q}{\partial z_i^{\sigma}} = -\frac{\alpha_i \varphi_i}{\sum_{j=1}^{m} \alpha_j \varphi_j}\left\{\frac{\|t - \mu_i\|^2}{\sigma_i^2} - c\right\}$$

- Mixture Density LSTM

1 $$p(\boldsymbol{t}|\boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i(\boldsymbol{x})\varphi_i(\boldsymbol{t}|\boldsymbol{x})$$

2 $$\varphi_i(\boldsymbol{t}|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{c}{2}}\sigma_i(\boldsymbol{x})^c} e^{-\frac{\|\boldsymbol{t}-\boldsymbol{\mu}_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}}$$

3 $$\boldsymbol{z} = [z_1^\alpha, \ldots, z_m^\alpha, z_{m+1}^\mu, \ldots, z_{mc+m+1}^\mu, z_{mc+m+2}^\sigma, \ldots, z_{m(c+2)}^\sigma]$$

4 $$E^q = -\log\left[\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\varphi_i(\boldsymbol{t}^q|\boldsymbol{x}^q)\right]$$

5 $$\alpha_i = \frac{e^{z_i^\alpha}}{\sum_{j=1}^{M} e^{z_j^\alpha}}, \quad \sigma_i = e^{z_i^\sigma}, \quad \mu_i = z_{ik}^\mu$$

$$\frac{\partial E^q}{\partial z_k^\alpha} = \alpha_k - \frac{\alpha_k\varphi_k}{\sum_{j=1}^{m} \alpha_j\varphi_j}$$

$$\frac{\partial E^q}{\partial z_{ik}^\mu} = \frac{\alpha_i\varphi_i}{\sum_{j=1}^{m} \alpha_j\varphi_j} \frac{\mu_{ik} - t_k}{\sigma_i^2}$$ 6

$$\frac{\partial E^q}{\partial z_i^\sigma} = -\frac{\alpha_i\varphi_i}{\sum_{j=1}^{m} \alpha_j\varphi_j}\left\{\frac{\|t-\mu_i\|^2}{\sigma_i^2} - c\right\}$$

Generative

- Mixture Density LSTM

1 $$p(\boldsymbol{t}|\boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i(\boldsymbol{x})\varphi_i(\boldsymbol{t}|\boldsymbol{x})$$

2 $$\varphi_i(\boldsymbol{t}|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{c}{2}}\sigma_i(\boldsymbol{x})^c} e^{-\frac{\|\boldsymbol{t}-\boldsymbol{\mu}_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}}$$

3 $$\boldsymbol{z} = [z_1^{\alpha}, \ldots, z_m^{\alpha}, z_{m+1}^{\mu}, \ldots, z_{mc+m+1}^{\mu}, z_{mc+m+2}^{\sigma}, \ldots, z_{m(c+2)}^{\sigma}]$$

4 $$E^q = -\log\left[\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\varphi_i(\boldsymbol{t}^q|\boldsymbol{x}^q)\right]$$

5 $$\alpha_i = \frac{e^{z_i^{\alpha}}}{\sum_{j=1}^{M} e^{z_j^{\alpha}}}, \quad \sigma_i = e^{z_i^{\sigma}}, \quad \mu_i = z_{ik}^{\mu}$$

$$\frac{\partial E^q}{\partial z_k^{\alpha}} = \alpha_k - \frac{\alpha_k \varphi_k}{\sum_{j=1}^{m} \alpha_j \varphi_j}$$

$$\frac{\partial E^q}{\partial z_{ik}^{\mu}} = \frac{\alpha_i \varphi_i}{\sum_{j=1}^{m} \alpha_j \varphi_j} \frac{\mu_{ik} - t_k}{\sigma_i^2}$$ 6

$$\frac{\partial E^q}{\partial z_i^{\sigma}} = -\frac{\alpha_i \varphi_i}{\sum_{j=1}^{m} \alpha_j \varphi_j} \left\{\frac{\|t - \mu_i\|^2}{\sigma_i^2} - c\right\}$$

Generative

- Mixture Density LSTM

1 $$p(\boldsymbol{t}|\boldsymbol{x}) = \sum_{i=1}^{m} \alpha_i(\boldsymbol{x})\varphi_i(\boldsymbol{t}|\boldsymbol{x})$$

2 $$\varphi_i(\boldsymbol{t}|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{c}{2}}\sigma_i(\boldsymbol{x})^c} e^{-\frac{\|\boldsymbol{t}-\boldsymbol{\mu}_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}}$$

3 $$\boldsymbol{z} = [z_1^\alpha, \ldots, z_m^\alpha, z_{m+1}^\mu, \ldots, z_{mc+m+1}^\mu, z_{mc+m+2}^\sigma, \ldots, z_{m(c+2)}^\sigma]$$

4 $$E^q = -\log\left[\sum_{i=1}^{m} \alpha_i(\boldsymbol{x}^q)\varphi_i(\boldsymbol{t}^q|\boldsymbol{x}^q)\right]$$

5 $$\alpha_i = \frac{e^{z_i^\alpha}}{\sum_{j=1}^{M} e^{z_j^\alpha}}, \quad \sigma_i = e^{z_i^\sigma}, \quad \mu_i = z_{ik}^\mu$$

$$\frac{\partial E^q}{\partial z_k^\alpha} = \alpha_k - \frac{\alpha_k\varphi_k}{\sum_{j=1}^{m}\alpha_j\varphi_j}$$

$$\frac{\partial E^q}{\partial z_{ik}^\mu} = \frac{\alpha_i\varphi_i}{\sum_{j=1}^{m}\alpha_j\varphi_j}\frac{\mu_{ik}-t_k}{\sigma_i^2}$$ 6

$$\frac{\partial E^q}{\partial z_i^\sigma} = -\frac{\alpha_i\varphi_i}{\sum_{j=1}^{m}\alpha_j\varphi_j}\left\{\frac{\|t-\mu_i\|^2}{\sigma_i^2} - c\right\}$$

Generative

- Mixture Density LSTM

$$\frac{\partial E^q}{\partial z_k^\alpha} = \alpha_k - \frac{\alpha_k \varphi_k}{\sum_{j=1}^m \alpha_j \varphi_j}$$

1 $$p(\boldsymbol{t}|\boldsymbol{x}) = \sum_{i=1}^m \alpha_i(\boldsymbol{x}) \varphi_i(\boldsymbol{t}|\boldsymbol{x})$$

$$\frac{\partial E^q}{\partial z_{ik}^\mu} = \frac{\alpha_i \varphi_i}{\sum_{j=1}^m \alpha_j \varphi_j} \frac{\mu_{ik} - t_k}{\sigma_i^2}$$ 6

2 $$\varphi_i(\boldsymbol{t}|\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{c}{2}} \sigma_i(\boldsymbol{x})^c} e^{-\frac{\|\boldsymbol{t} - \boldsymbol{\mu}_i(\boldsymbol{x})\|^2}{2\sigma_i(\boldsymbol{x})^2}}$$

$$\frac{\partial E^q}{\partial z_i^\sigma} = -\frac{\alpha_i \varphi_i}{\sum_{j=1}^m \alpha_j \varphi_j} \left\{ \frac{\|t - \mu_i\|^2}{\sigma_i^2} - c \right\}$$

3 $$\boldsymbol{z} = [z_1^\alpha, \ldots, z_m^\alpha, z_{m+1}^\mu, \ldots, z_{mc+m+1}^\mu, z_{mc+m+2}^\sigma, \ldots, z_{m(c+2)}^\sigma]$$

4 $$E^q = -\log \left[\sum_{i=1}^m \alpha_i(\boldsymbol{x}^q) \varphi_i(\boldsymbol{t}^q|\boldsymbol{x}^q) \right]$$

5 $$\alpha_i = \frac{e^{z_i^\alpha}}{\sum_{j=1}^M e^{z_j^\alpha}}, \quad \sigma_i = e^{z_i^\sigma}, \quad \mu_i = z_{ik}^\mu$$

Wanna be Doing Deep Learning?

Deep Learning with Python

python has a wide range of deep learning-related libraries available

and of course: scikit learn

High level

Keras — keras.io

Lasagne — lasagne.readthedocs.org/en/latest

TensorFlow — tensorflow.org/

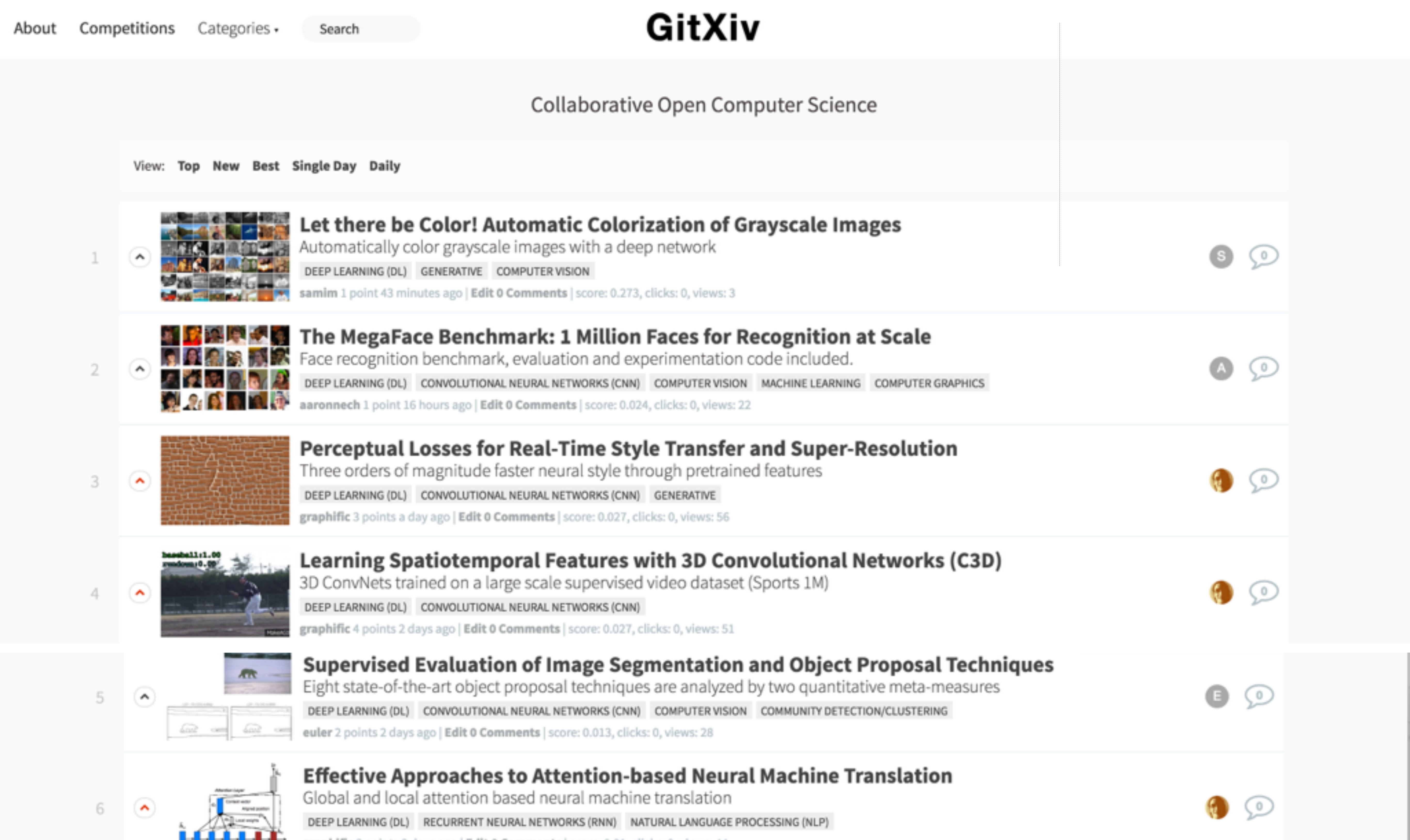
Caffe — caffe.berkeleyvision.org

Low level

theano — deeplearning.net/software/theano

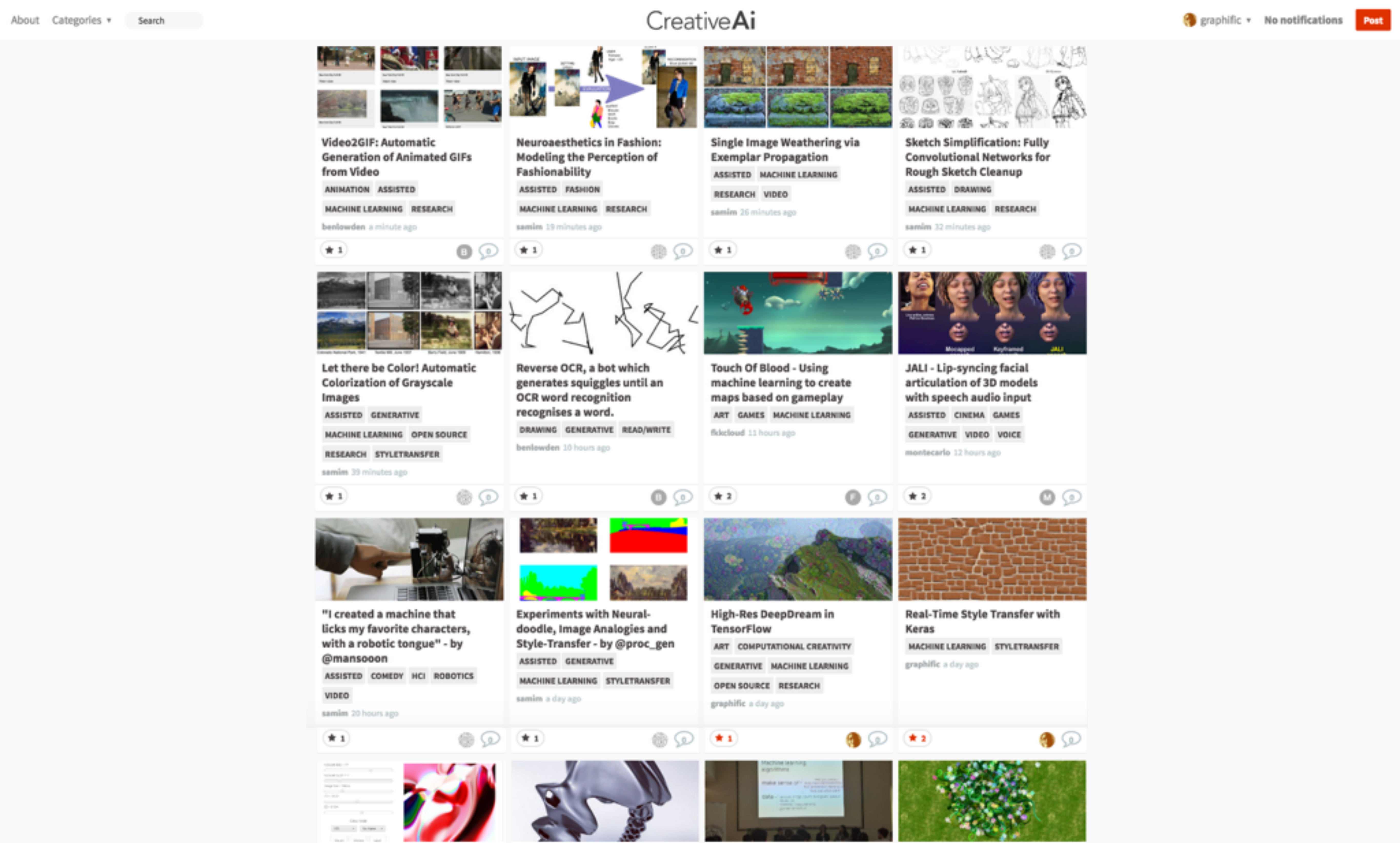
Code & Papers?

http://gitxiv.com/ #GitXiv



Creative AI projects?

http://www.creativeai.net/ #CreativeAI



Questions?

love letters? existential dilemma's? academic questions? gifts?

find me at:

www.csc.kth.se/~roelof/

roelof@kth.se

@graphific

KTH VETENSKAP OCH KONST

ROYAL INSTITUTE OF TECHNOLOGY

Oh, and soon we're looking for Creative AI enthusiasts !

AE ARTIFICIAL EXPERIENCE

- job
- internship
- thesis work

in **AI (Deep Learning) & Creativity**

Medium

https://medium.com/@ArtificialExperience/creativeai-9d4b2346faf3

CreativeAI

On the Democratisation & Escalation of Creativity

Chapter 01

by Samim Winiger & Roelof Pieters

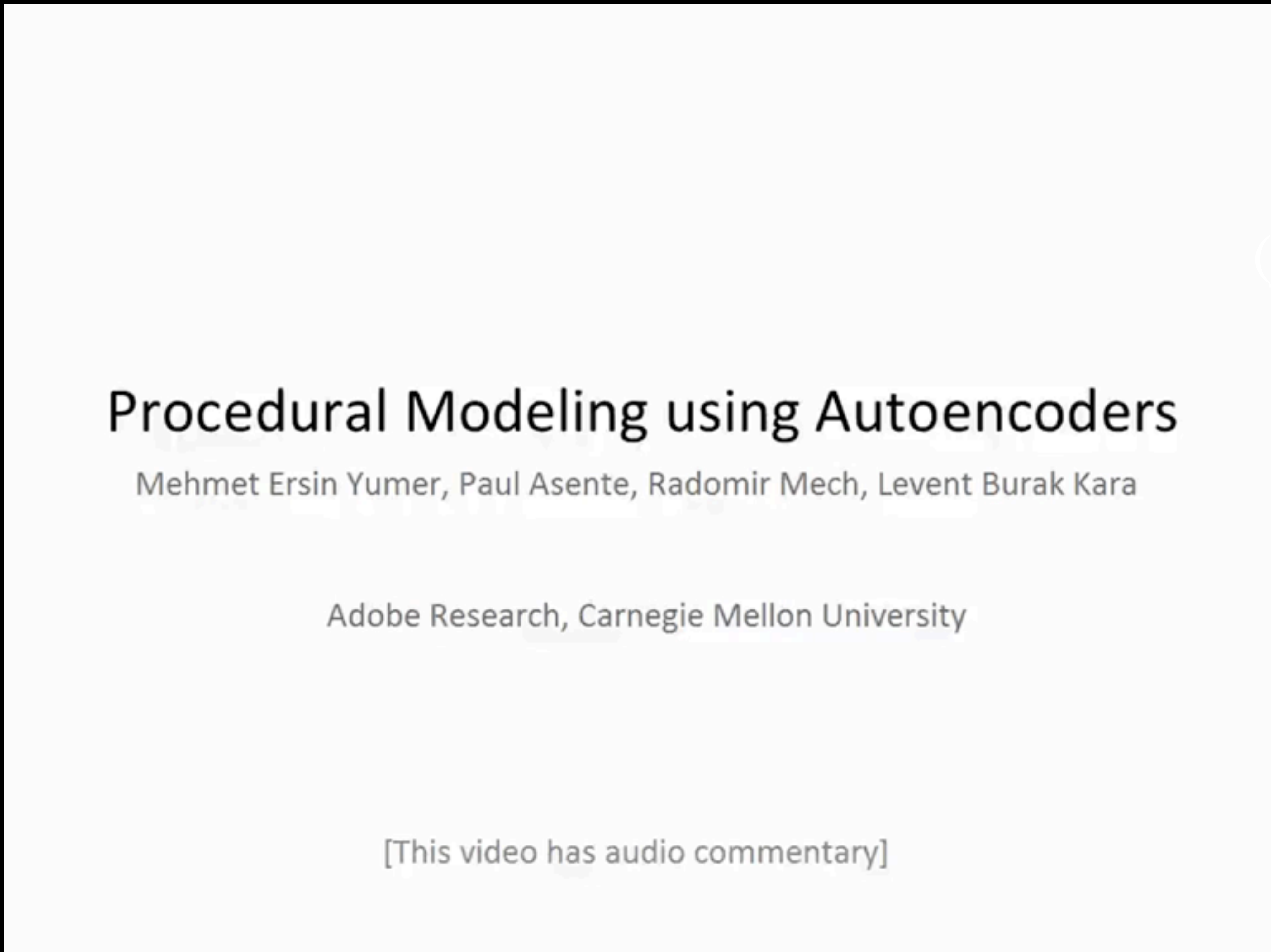
human-machine collaboration

Semantic Shape Editing Using Deformation Handles (YouTube, Paper)

M.E. Yumer, S. Chaudhuri, J.K. Hodgins, L.B. Kara

SIGGRAPH 2015

Narrated by Jennifer Kara



(YouTube, Paper)

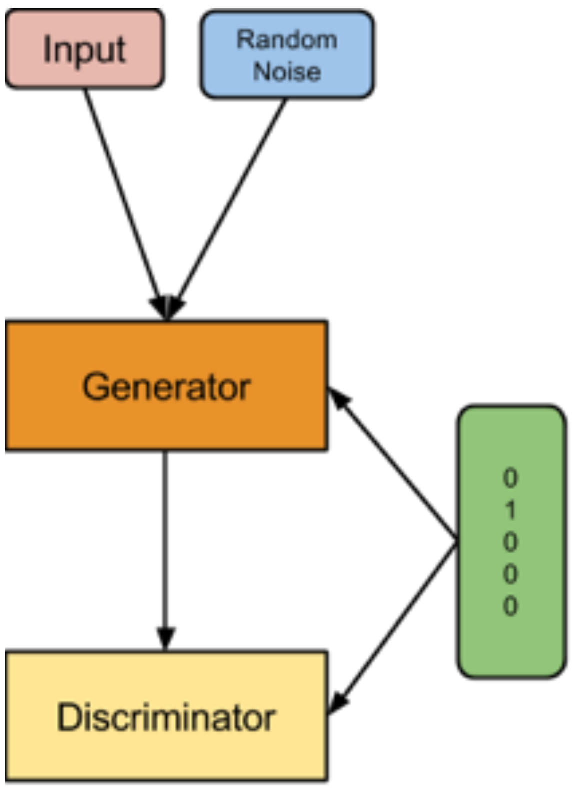
(Vimeo, Paper)

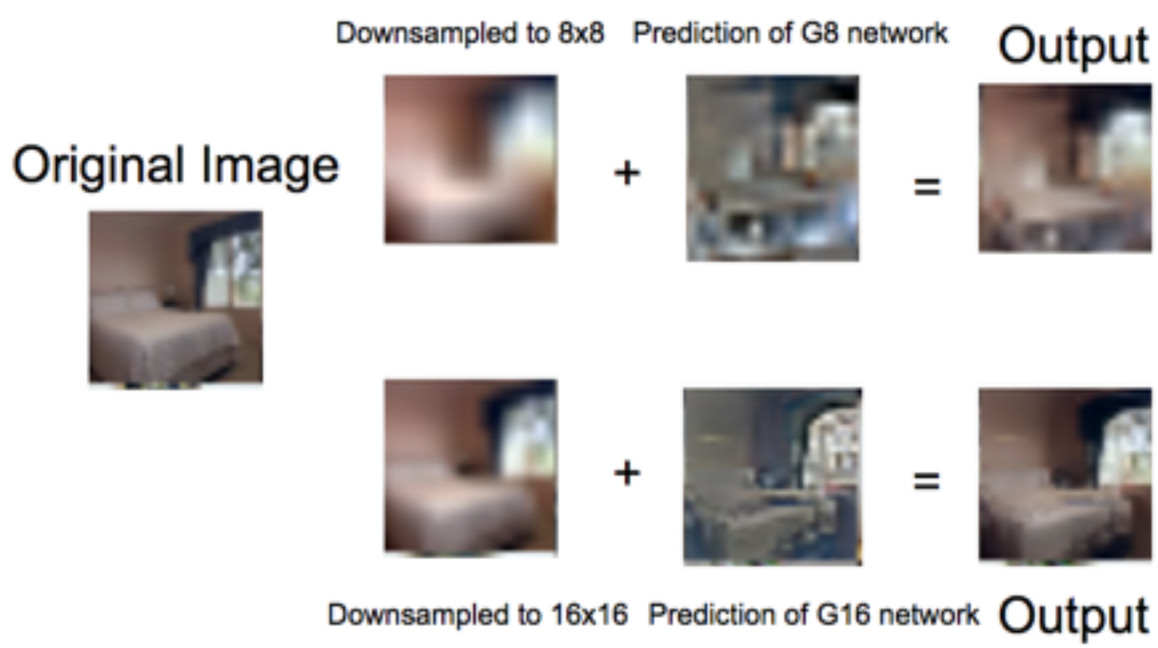
Exploratory Modeling with Collaborative Design Spaces

Jerry O. Talton   Daniel Gibson   Lingfeng Yang
Pat Hanrahan   Vladlen Koltun

Stanford University

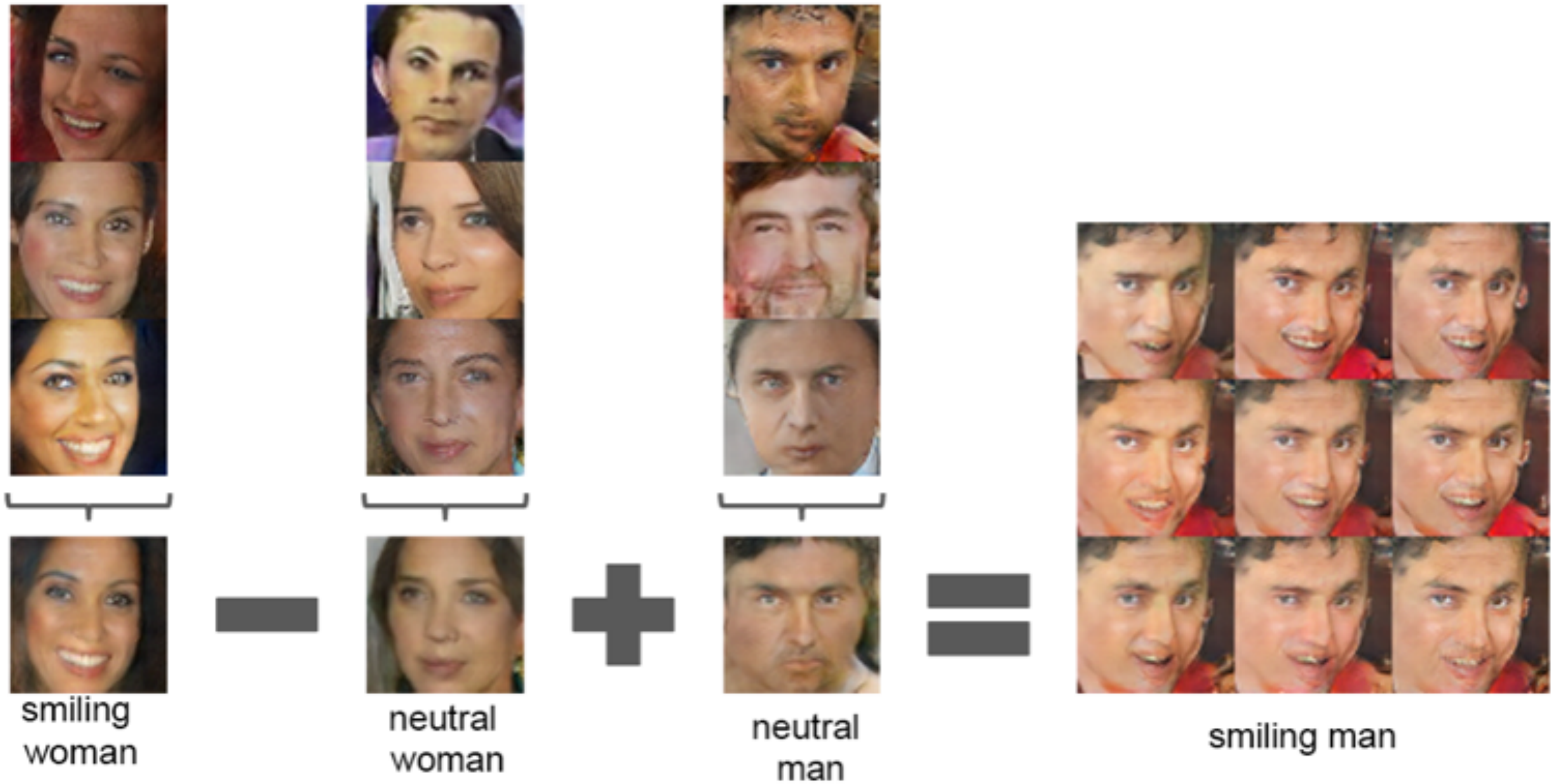
Generative Adverserial Nets





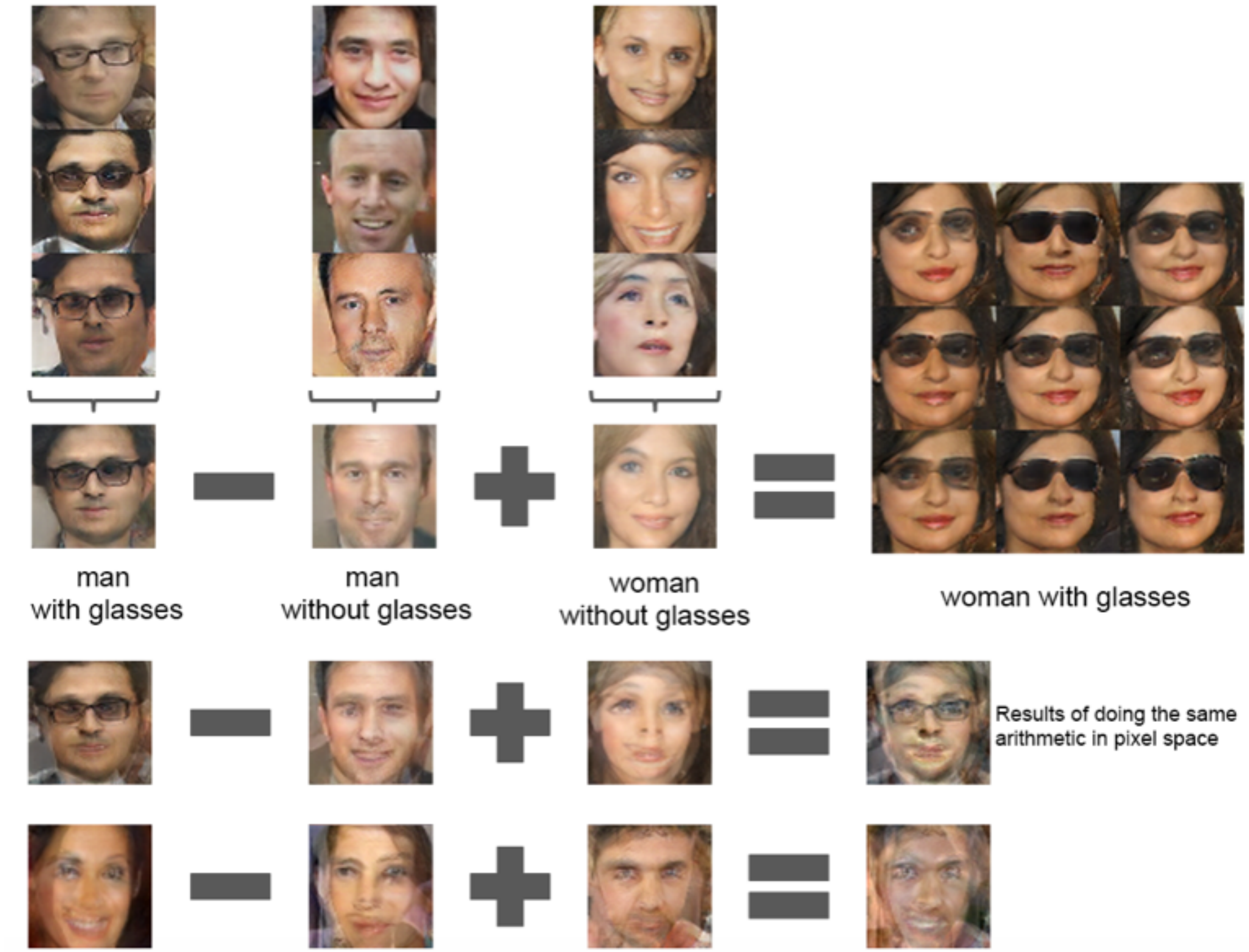
Emily Denton, Soumith Chintala, Arthur Szlam, Rob Fergus, 2015.
Deep Generative Image Models using a Laplacian Pyramid of Adversarial Networks (GitXiv)

Generative Adverserial Nets



Alec Radford, Luke Metz, Soumith Chintala , 2015.
Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (GitXiv)

Generative Adverserial Nets



Alec Radford, Luke Metz, Soumith Chintala , 2015.
Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (GitXiv)

Generative Adverserial Nets

”turn” vector created from four averaged samples of faces looking left vs looking right.

Alec Radford, Luke Metz, Soumith Chintala , 2015.
Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks (GitXiv)

Generative Adverserial Nets

walking through the manifold

Generative Adverserial Nets

top: unmodified samples
bottom: same samples dropping out ”window” filters