

The background of the slide is a photograph of a large, multi-story library. A prominent white spiral staircase with a dark metal railing winds through the center of the space. The walls and balconies are lined with tall, dark wood bookshelves filled with books. The lighting is warm and even, highlighting the architectural details and the vast collection of literature.

курс «Глубокое обучение»

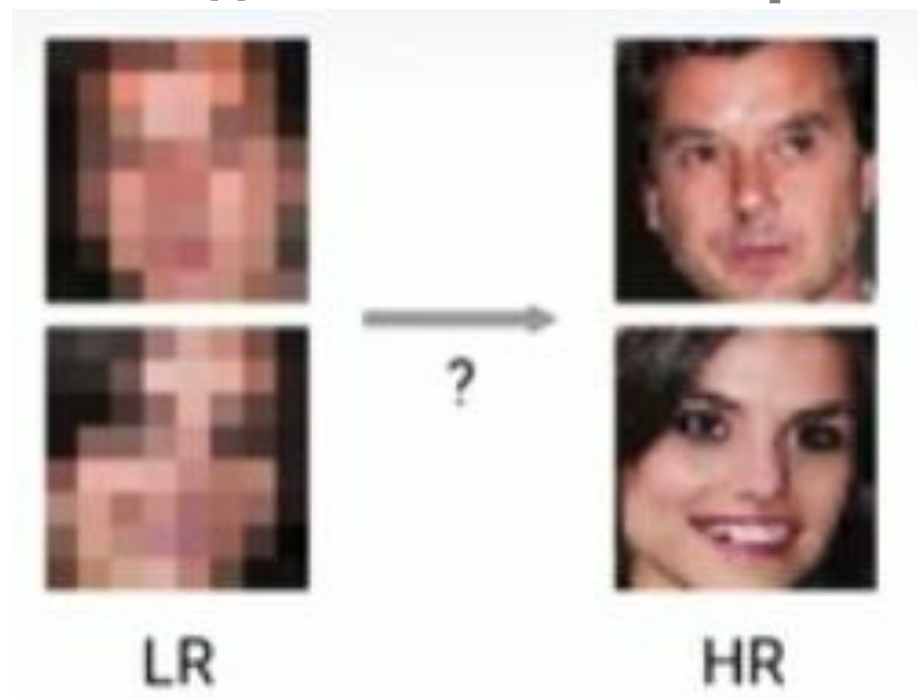
CNN: дополнение

Александр Дьяконов
(ВМК МГУ имени М.В. Ломоносова)

23 марта 2020 года

Deep image prior (DIP)

Prior – наши знания о мире
«как выглядят естественные решения»

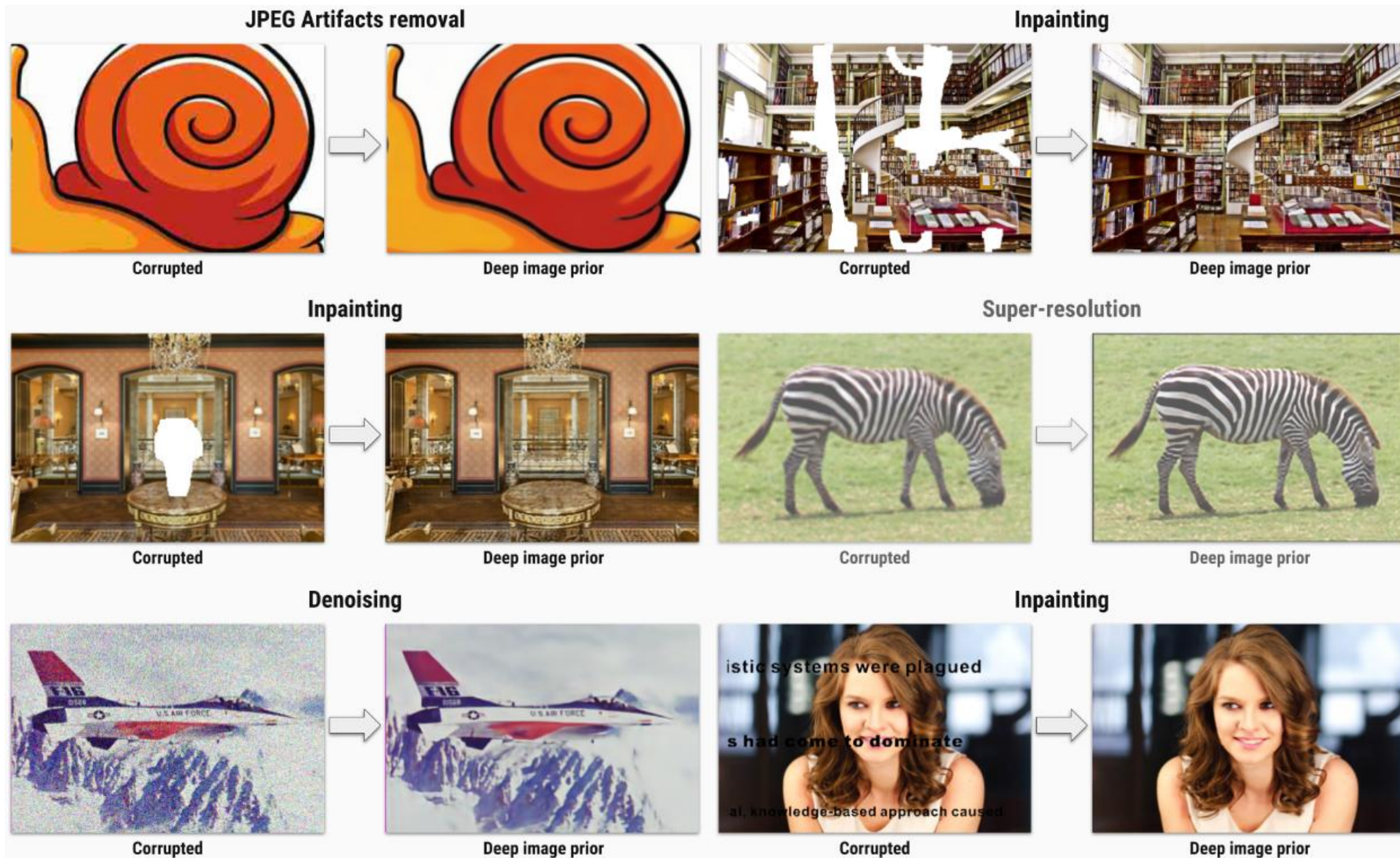


возможно отображение
изображение → изображение

- по одному изображению
 - робастно
 - без учителя
- без предтренировки

<https://www.youtube.com/watch?v=FyUhnjuOEXo>

Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9446–9454, 2018.



Задача восстановления изображения – Image Restoration

Обычно задача восстановления:

$$\min_x E(x, x_0) + R(x)$$

x_0 – повреждённое изображение

$E(x, x_0)$ – легко придумать функцию «под задачу»

$R(x)$ – «image prior» – как придумать?!

чтобы описывались «естественные изображения»

Почему так...

$$x^* = \arg \max_x p(x | x_0)$$

$$p(x | x_0) = p(x_0 | x) p(x) / p(x_0) \propto \underbrace{p(x_0 | x)}_{\text{Likelihood}} \underbrace{p(x)}_{\text{Prior}}$$

$$-\log p(x | x_0) \propto -\log p(x_0 | x) - \log p(x)$$

Модель зашумления



если есть модель зашумления

$$x_0 = x + \varepsilon, \varepsilon \sim \text{norm}(0, \sigma)$$
$$p(x_0 | x) = \text{norm}(x_0; x, \sigma)$$

Deep image prior (DIP): идея

$$\min_x E(x, x_0) + R(x)$$

пусть $x = g_\theta(z)$

g_θ – **CNN**

z – **фиксированный вход (не важно какой!)**

$$\min_\theta E(g_\theta(z), x_0) + R(g_\theta(z))$$

**и тут регуляризация не в пространстве изображений,
а параметров**

пусть регуляризация – возможность реализации...

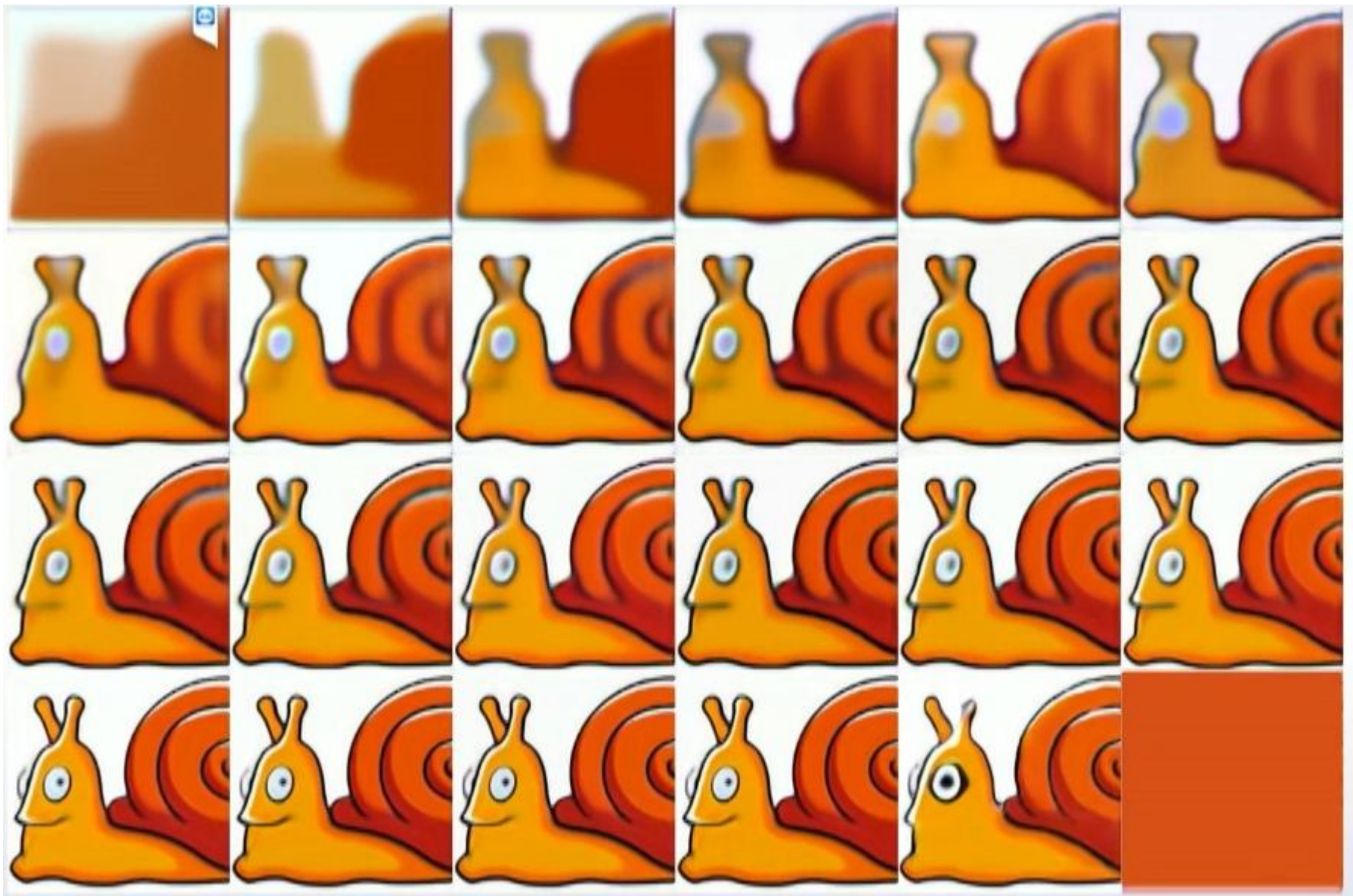
$$R(x) = +\infty \cdot I[\exists z : x = g_\theta(z)]$$

$$\min_\theta E(g_\theta(z), x_0)$$

prior ~ выбор архитектуры сети!

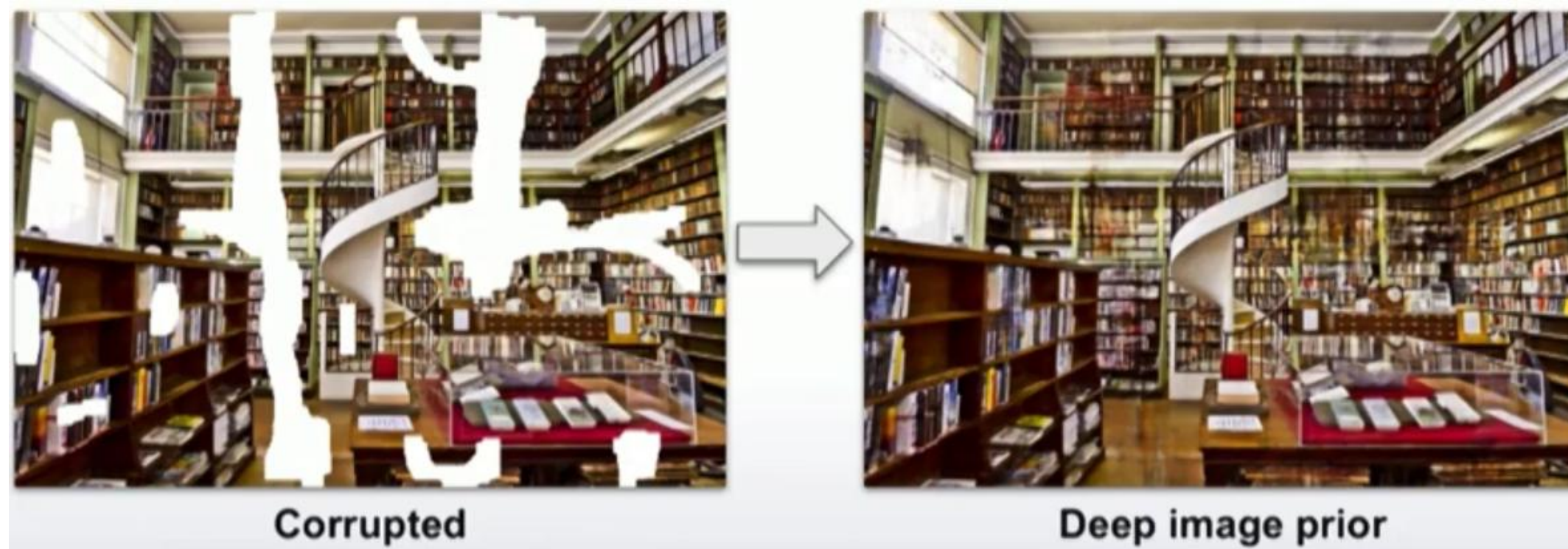
интересно, что нужен ранний останов

Deep image prior (DIP): Image Restoration



после некоторой итерации становится всё плохо!

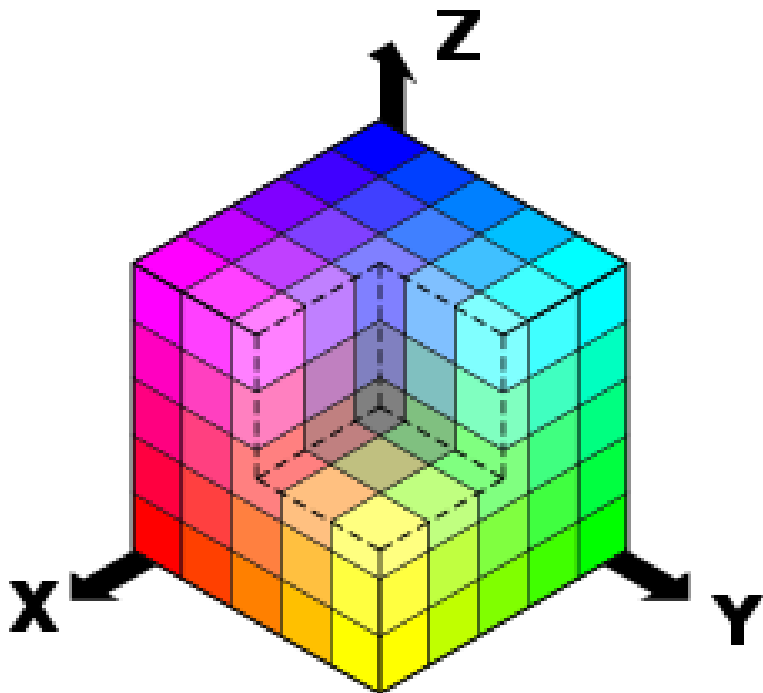
Deep image prior (DIP): Image Restoration



Раскраска изображений

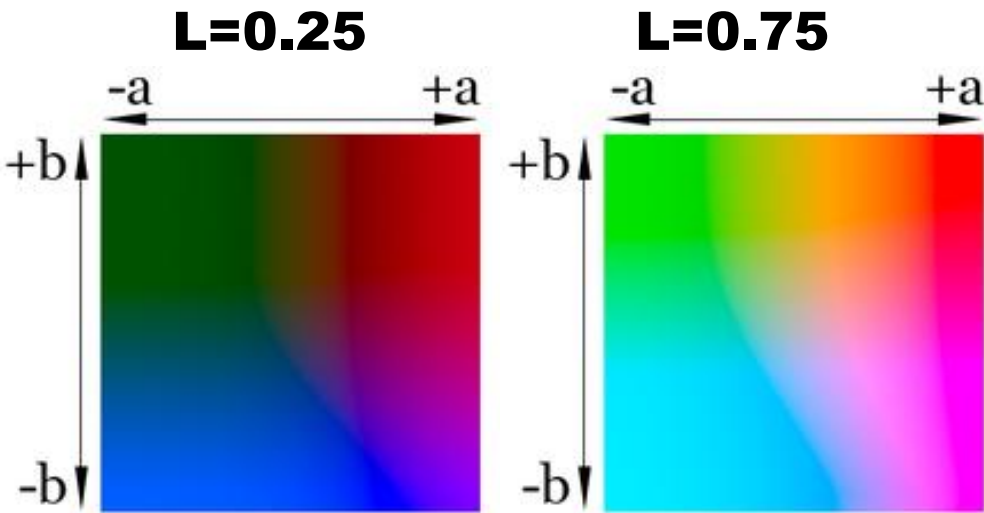
RGB

R: Red
G: Green
B: Blue



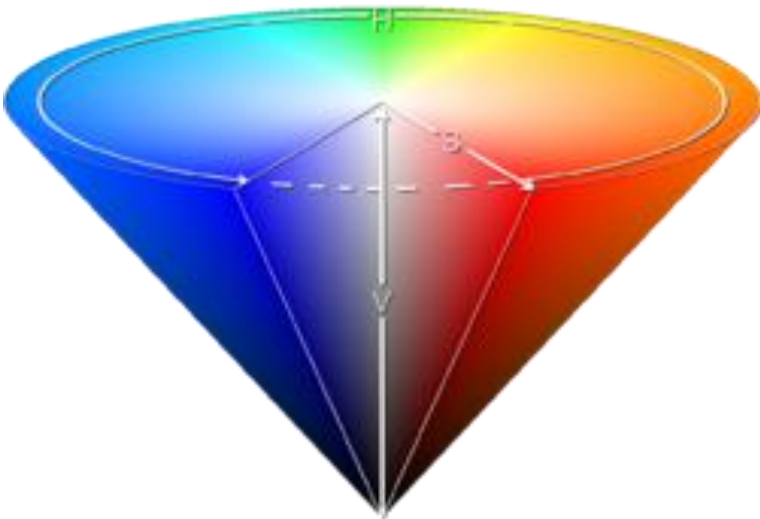
Lab

L: White-Black
a: Red-Green
b: Blue-Yellow



HSL / HSV

C: Chroma
H: Hue
L: White-Black



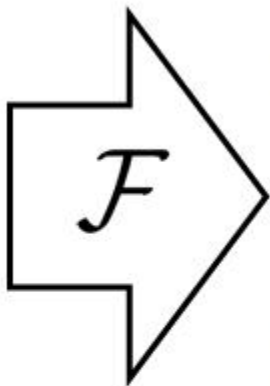
Раскраска изображений



Lab – «градации серого», «координаты цвета в шкале зелёный-красный», «координаты цвета в шкале синий-жёлтый»
осталось вычислить 2 канала...

<https://habr.com/ru/company/nix/blog/342388/>

Раскраска изображений

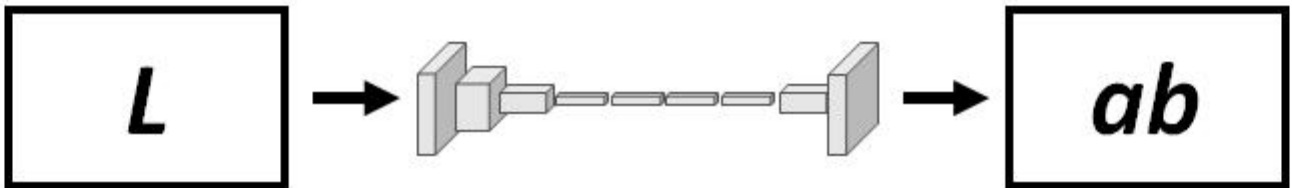


Grayscale image: L channel

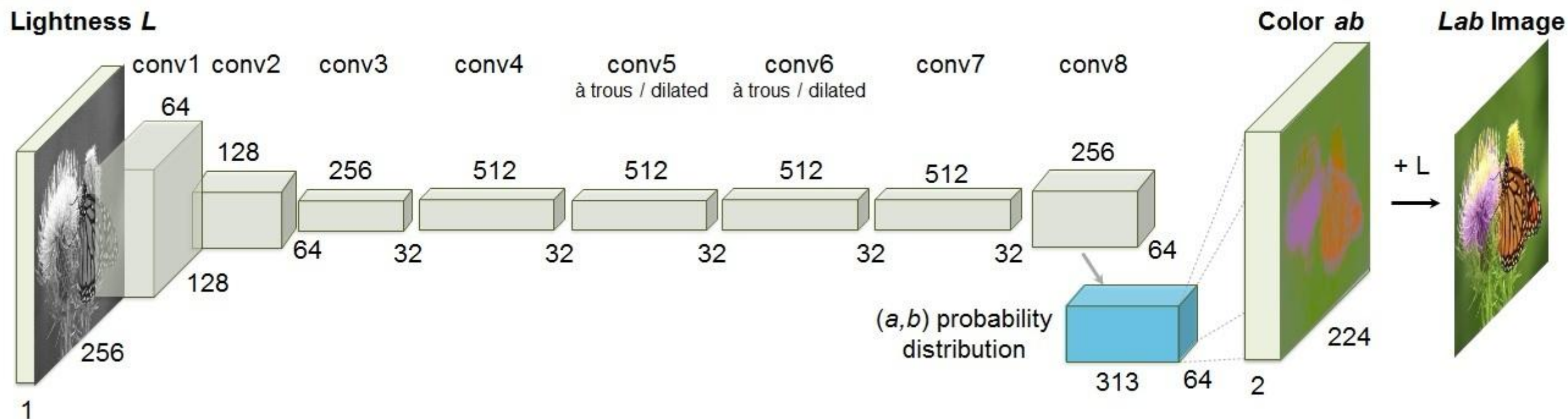
$$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$$

Color information: ab channels

$$\hat{\mathbf{Y}} \in \mathbb{R}^{H \times W \times 2}$$



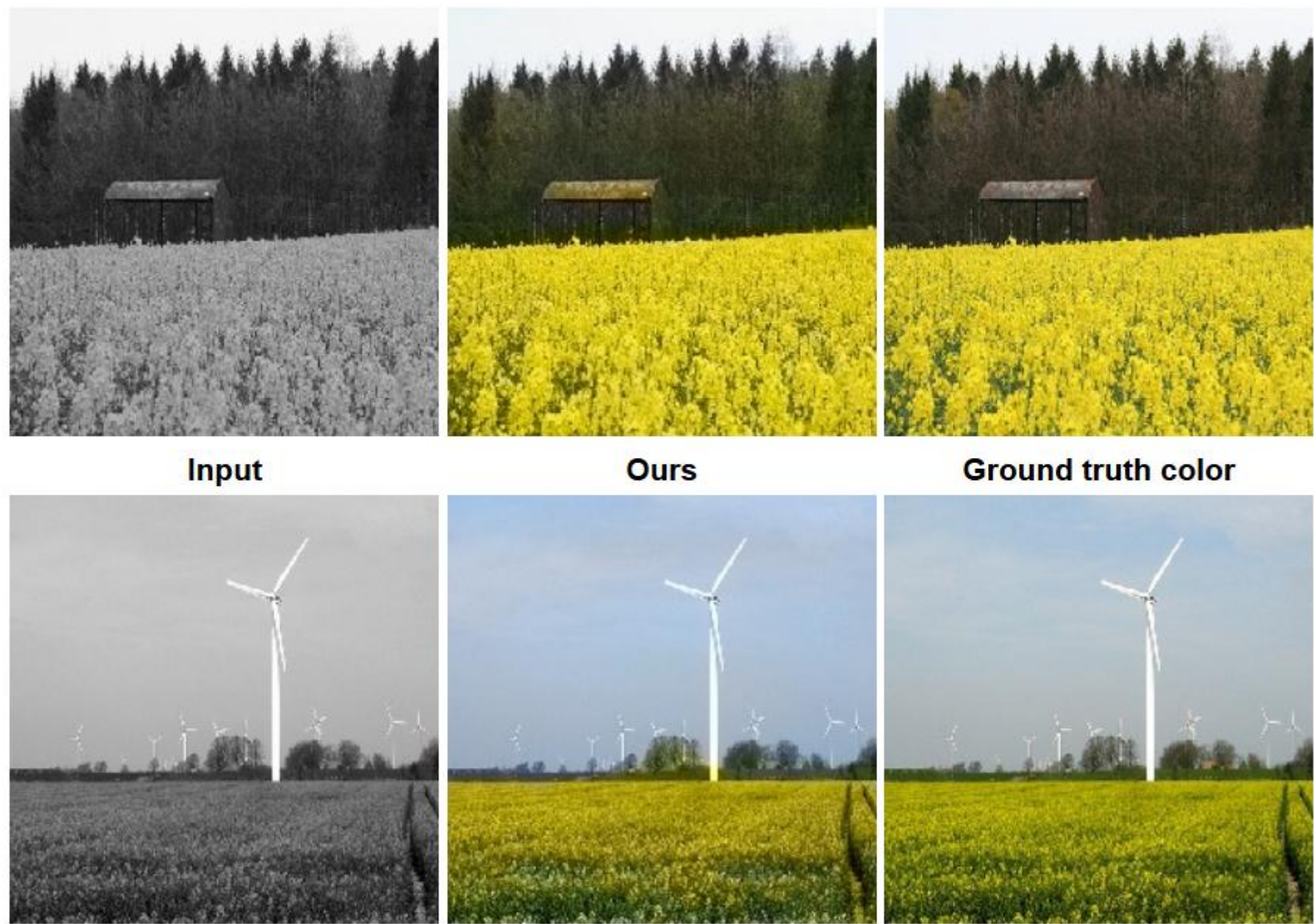
Раскраска изображений



каждый блок = 2 to 3 conv + ReLu-слои, за ними BN
нет пулинга
Dilated Convolutions

Richard Zhang, Phillip Isola, Alexei A. Efros «Colorful Image Colorization»
<https://richzhang.github.io/colorization/>

Раскраска изображений



Раскраска изображений: тонкости с ф-ей ошибки

L2-ошибка неадекватна

трюки с прогнозированием редких цветов

Better Loss Function

- Regression with L2 loss inadequate

$$L_2(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \|\mathbf{Y}_{h,w} - \hat{\mathbf{Y}}_{h,w}\|_2^2$$

- Use **multinomial classification**

$$L(\hat{\mathbf{Z}}, \mathbf{Z}) = -\frac{1}{HW} \sum_{h,w} \sum_q \mathbf{Z}_{h,w,q} \log(\hat{\mathbf{Z}}_{h,w,q})$$

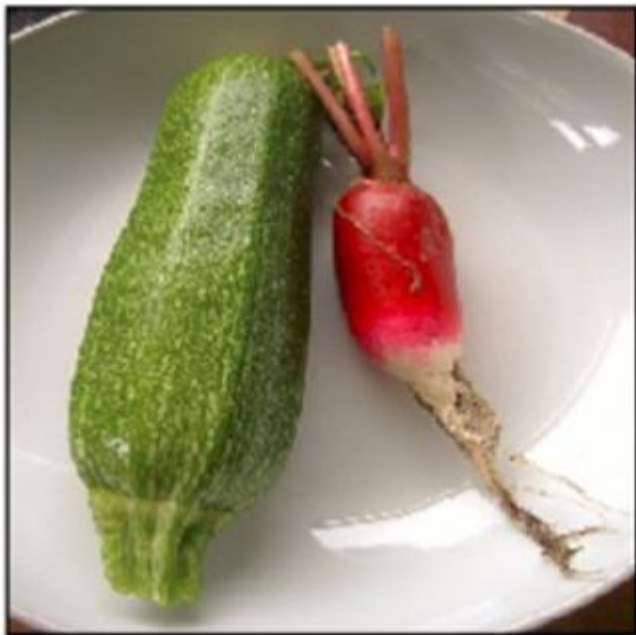
- **Class rebalancing** to encourage learning of *rare* colors

$$L(\hat{\mathbf{Z}}, \mathbf{Z}) = -\frac{1}{HW} \sum_{h,w} v(\mathbf{Z}_{h,w}) \sum_q \mathbf{Z}_{h,w,q} \log(\hat{\mathbf{Z}}_{h,w,q})$$

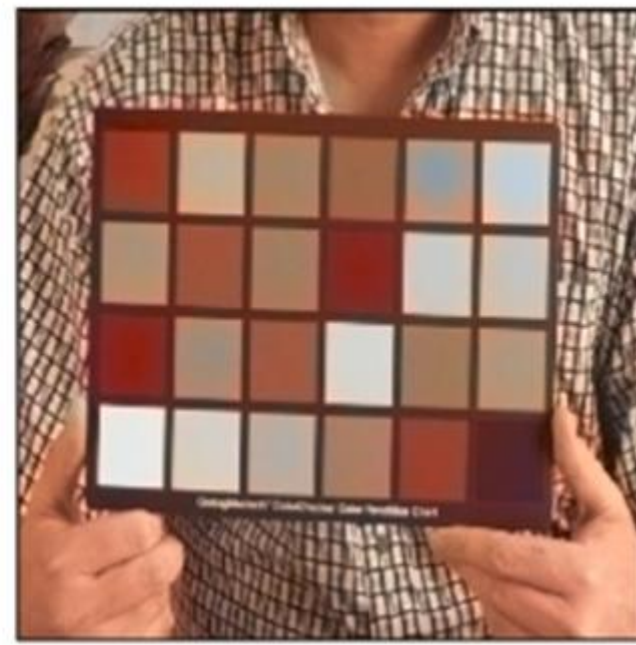
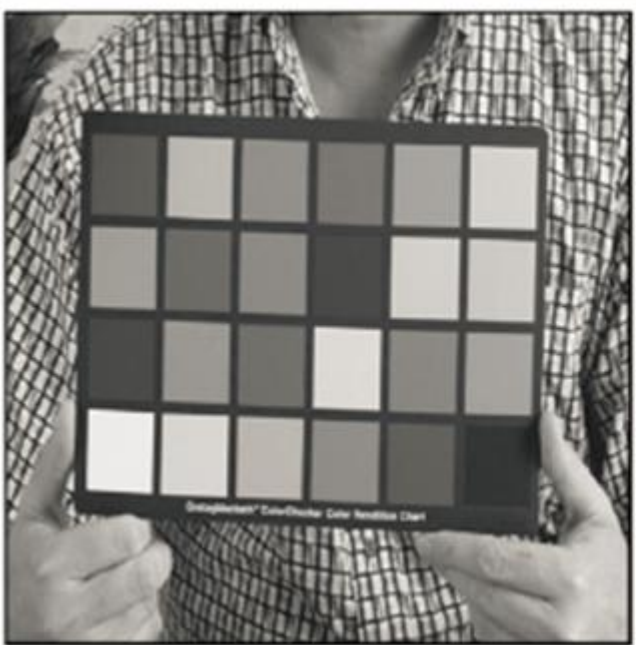
Input



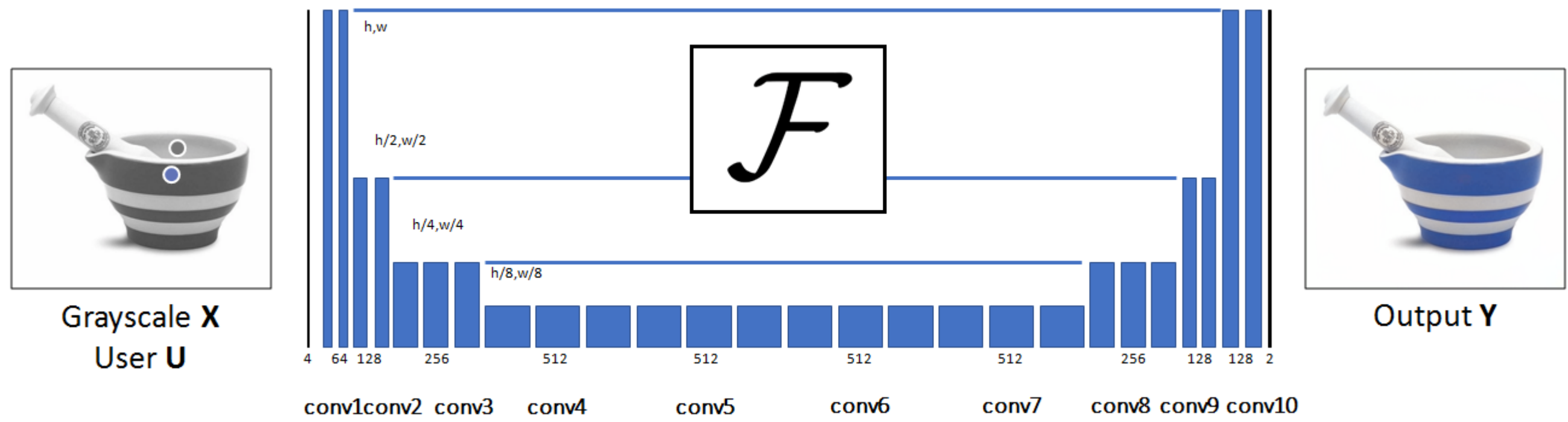
Ground Truth



Output



Раскраска изображений: следующая НС-архитектура авторов



но тут задача с разметкой цветов
<https://richzhang.github.io/ideepcolor/>

Раскраска с предварительной разметкой



Amateur family photograph, 1949.



Amateur family photograph, 1949.

Раскраска изображений

**Предсказание не отдельного цвета, а гистограммы
некоторые объекты (автомобили) имеют много допустимых цветов**

по окрестности $s \times s$ предсказываем цвет

есть хитрости с ф-ей ошибки... KL-дивергенция по распределениям цветов

Gustav Larsson, Michael Maire, Gregory Shakhnarovich «Learning Representations for Automatic Colorization» // <https://arxiv.org/abs/1603.06668>

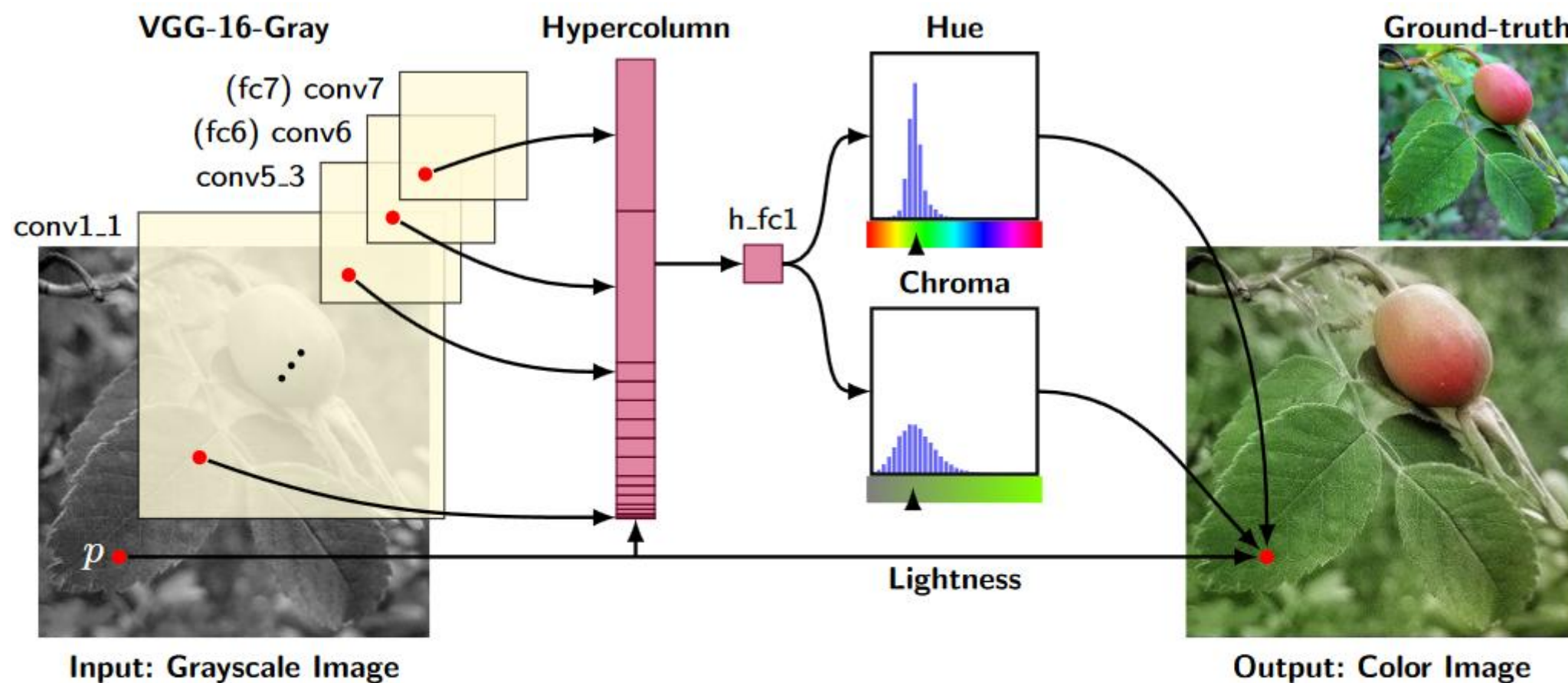


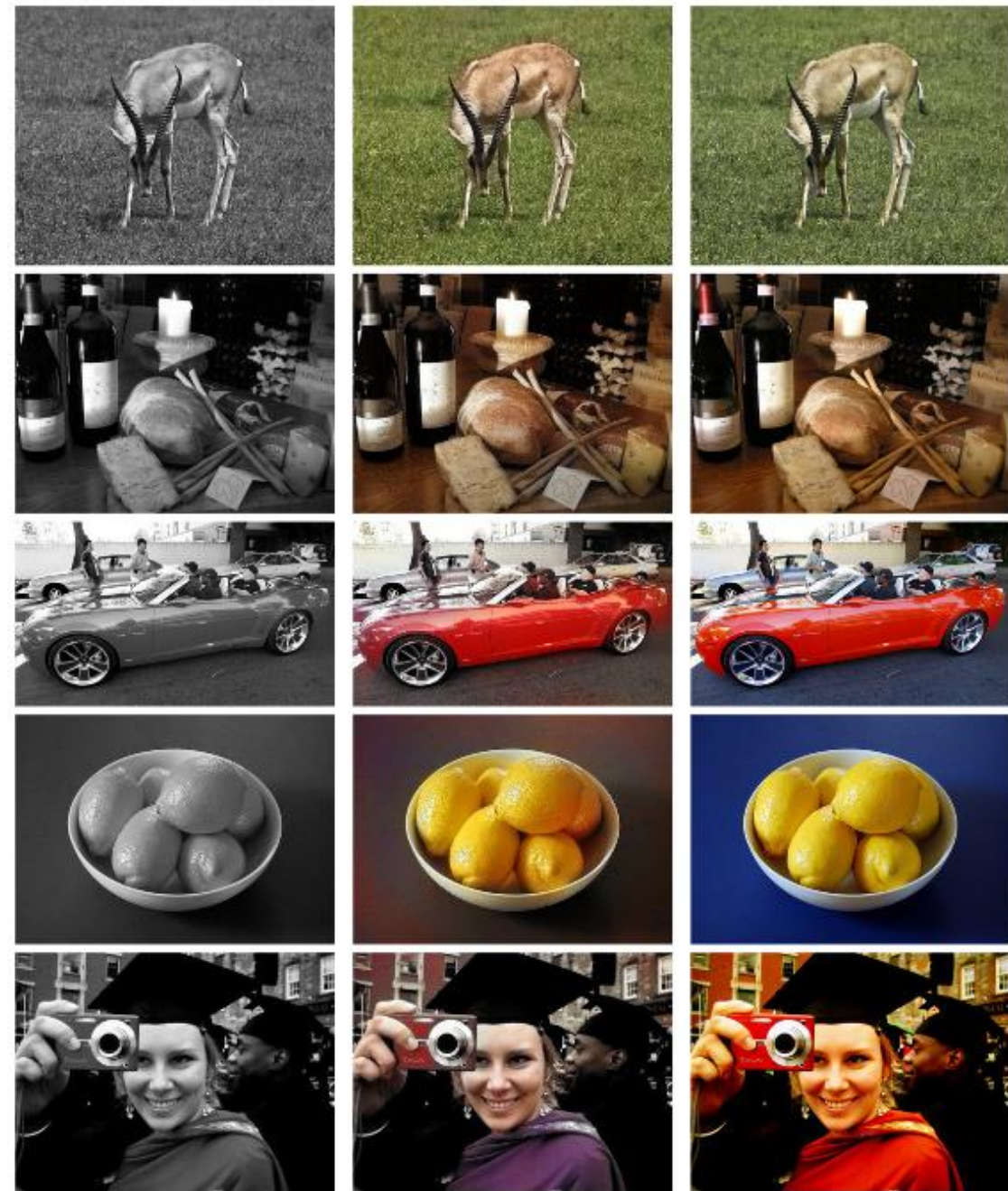
Fig. 2: **System overview.** We process a grayscale image through a deep convolutional architecture (VGG) [37] and take spatially localized multilayer slices (hypercolumns) [15, 26, 28], as per-pixel descriptors. We train our system end-to-end for the task of predicting hue and chroma distributions for each pixel p given its hypercolumn descriptor. These predicted distributions determine color assignment at test time.



Input

Our Method

Ground-truth



Input

Our Method

Ground-truth

Итог

в самой сети содержится мощный prior

**легко раскрашивать изображения
в первых методах были проблемы с функцией ошибки**