

neural image codec

Гарри Прошян M4150

Модель

Энкодер - все сверточные слои resnet18

Декодер - последовательность residual блоков вида:

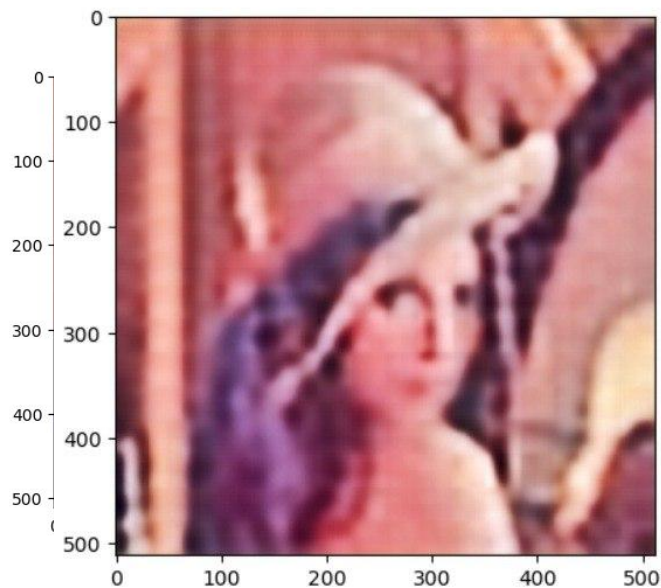
```
def forward(self, x):  
    skip_connection = self.conv_to_match_dims(x)  
    out = self.upscale(x)  
    out = self.double_conv(out)  
    out = out + skip_connection  
    return out
```

где double_conv = (Conv2d -> Batchnorm -> Relu) x2

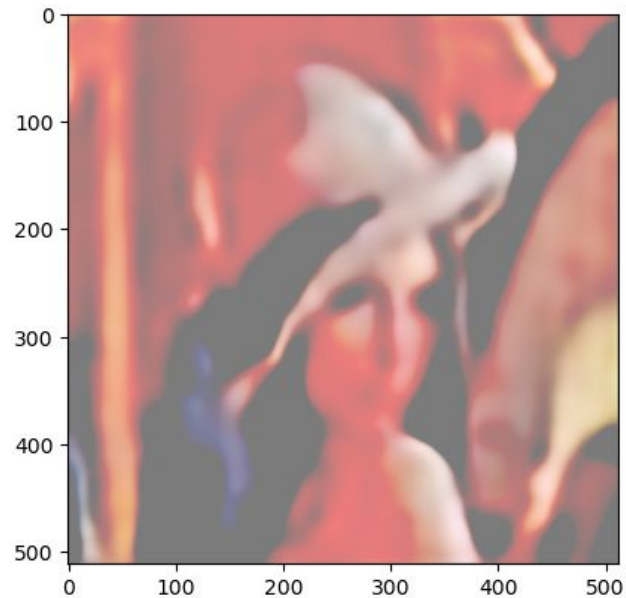
upscale = upsample

conv_to_match_dims - transpose convolution

Прирост качества при residual блоках



residual decoder



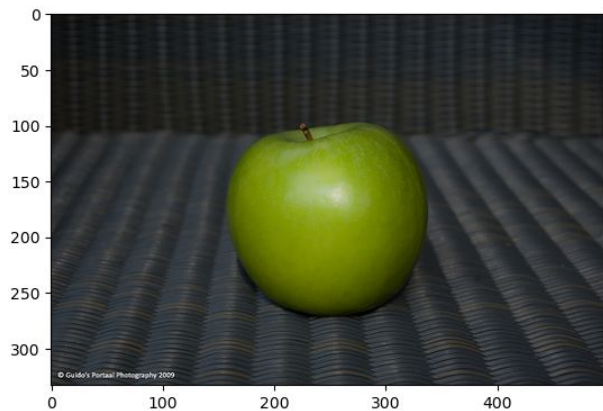
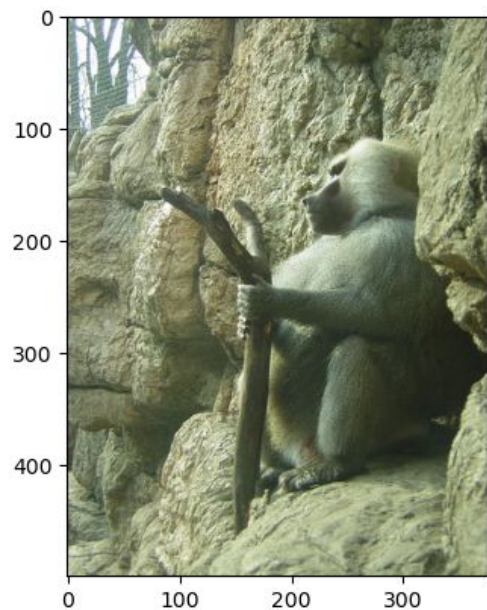
sequential decoder

Датасет

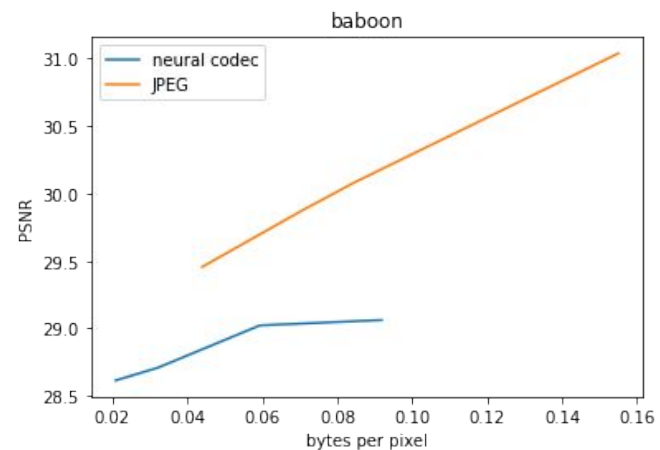
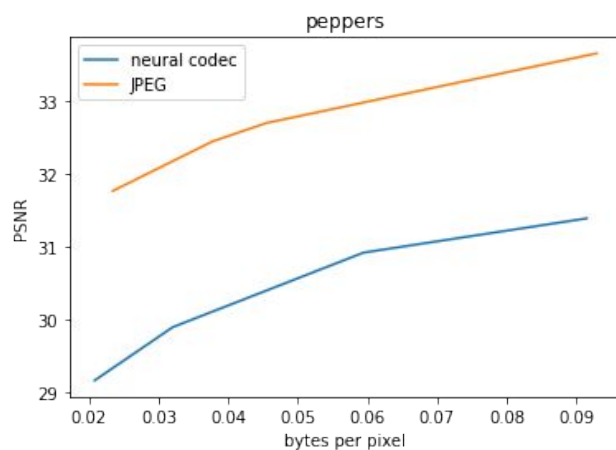
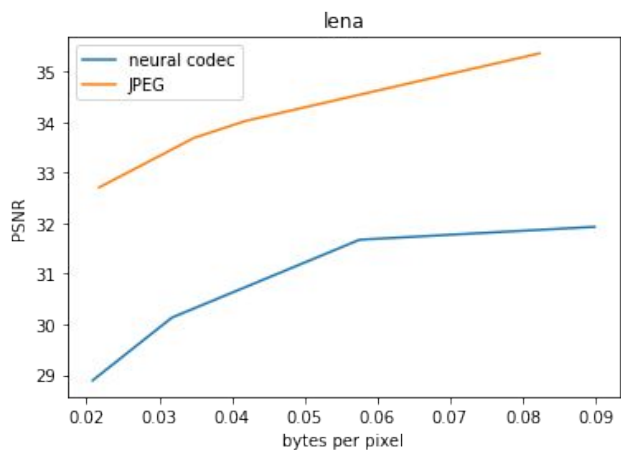
17'000 изображений из 13 классов из Imagenet:

```
"n02486410": "baboon",  
"n07720875": "bell pepper",  
"n03124170": "cowboy hat, ten-gallon hat",  
  
"n02493509": "titi, titi monkey",  
"n02110627": "affenpinscher, monkey pinscher, monkey dog",  
"n02493793": "spider monkey, Ateles geoffroyi",  
"n02480855": "gorilla, Gorilla gorilla",  
  
"n07742313": "Granny Smith",  
"n03724870": "mask",  
"n03379051": "football helmet",  
"n04356056": "sunglasses, dark glasses, shades",  
"n04591157": "Windsor tie",  
"n02906734": "broom"
```

Датасет

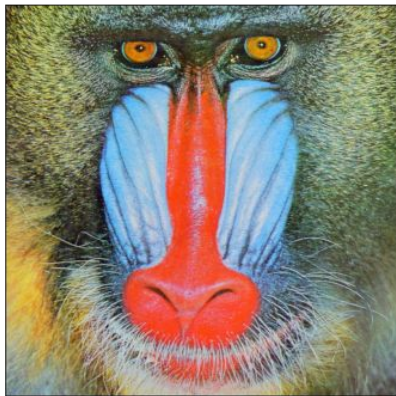


Результаты



$B = 6$

original



neural_codec

