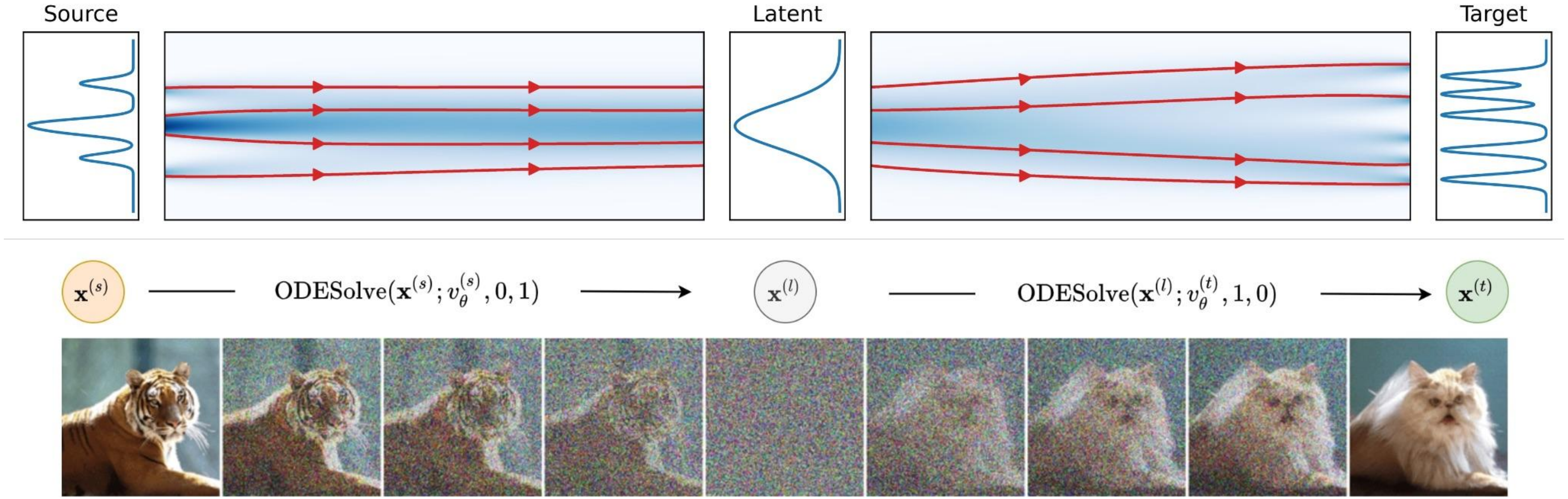
DDPM Image to Image

Alexander Sharshavin
Elfat Sabitov

Skoltech, Bayesian Methods

Problem statement

- Generate image from one domain by using style of the other image from the second
- Style transfer GANs, Normalising Flows, Diffusion Models



DDIB principle from [1]

DDIB

Solve probability flow ODE

$$\mathrm{d}\mathbf{x} = \left[\mathbf{f}(\mathbf{x}, t) - \frac{1}{2} g(t)^2 \nabla_{\mathbf{x}} \log p_t(\mathbf{x})\right] \mathrm{d}t$$

where score is approximated by a model: $\mathbf{s}_{t,\theta} \approx \nabla_{\mathbf{x}} \log p_t(\mathbf{x})$

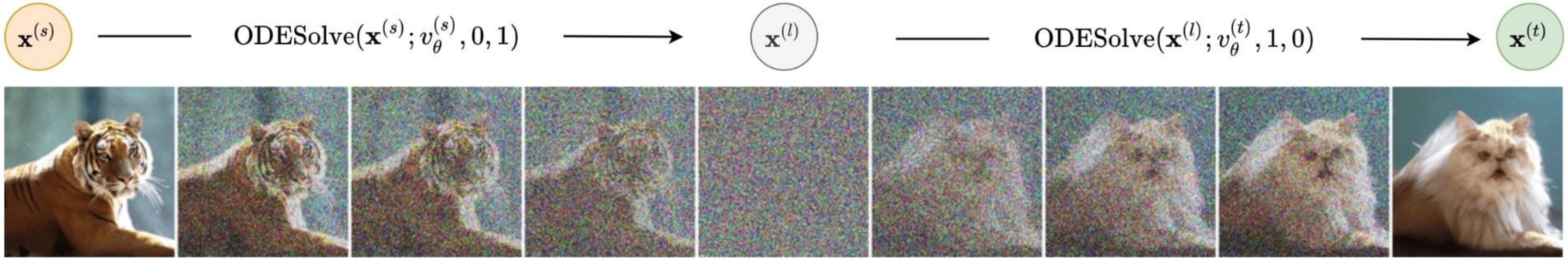
Thus, we deterministically solve an ODE:

$$\mathrm{ODESolve}(\mathbf{x}(t_0); v_\theta, t_0, t_1) = \mathbf{x}(t_0) + \int_{t_0}^{t_1} v_\theta(t, \mathbf{x}(t))\, \mathrm{d}t$$

DDIB

Algorithm 1 High-level Pseudo-code for DDIBs

Input: data sample from source domain $\mathbf{x}^{(s)} \sim p_s(\mathbf{x})$, source model $v_\theta^{(s)}$, target model $v_\theta^{(t)}$.
Output: $\mathbf{x}^{(t)}$, the result in the target domain.
$\mathbf{x}^{(l)} = \mathrm{ODESolve}(\mathbf{x}^{(s)}; v_\theta^{(s)}, 0, 1)$ // obtain latent code from source domain data
$\mathbf{x}^{(t)} = \mathrm{ODESolve}(\mathbf{x}^{(l)}; v_\theta^{(t)}, 1, 0)$ // obtain target domain data from latent code
return $\mathbf{x}^{(t)}$

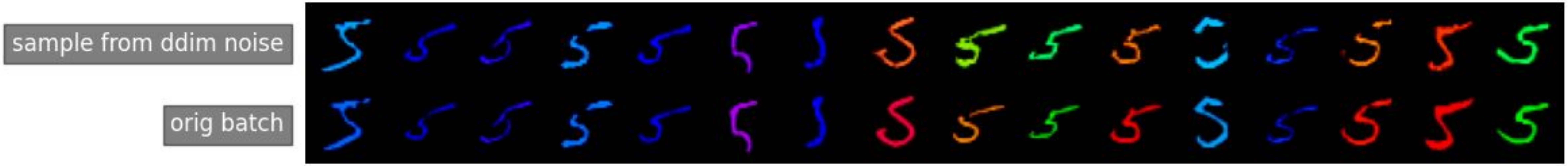


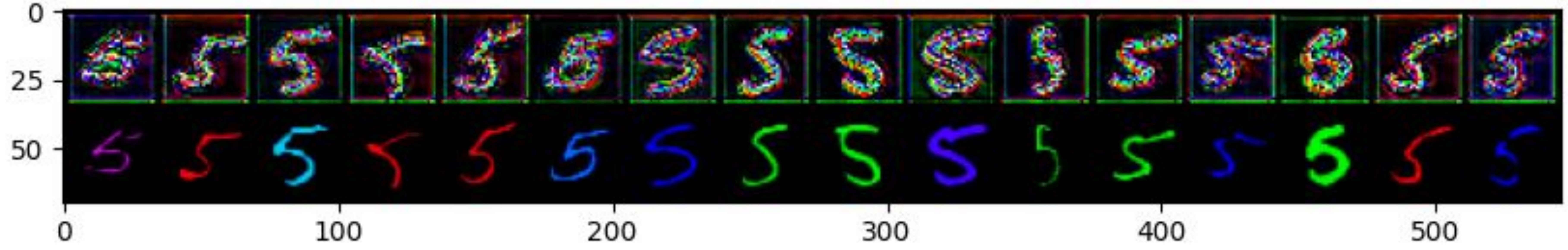
DDIB principle from [1]

Experiments

Data from colored MNST [2]

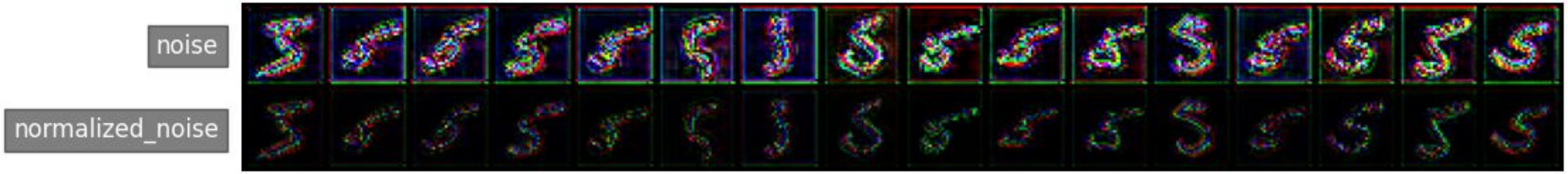
Try to generate digits 7 from this latent noise



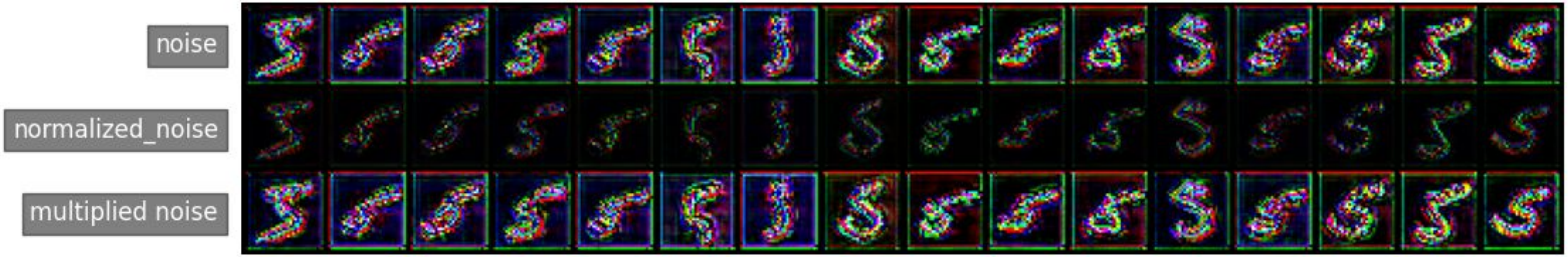


Experiments

Normalize noise:

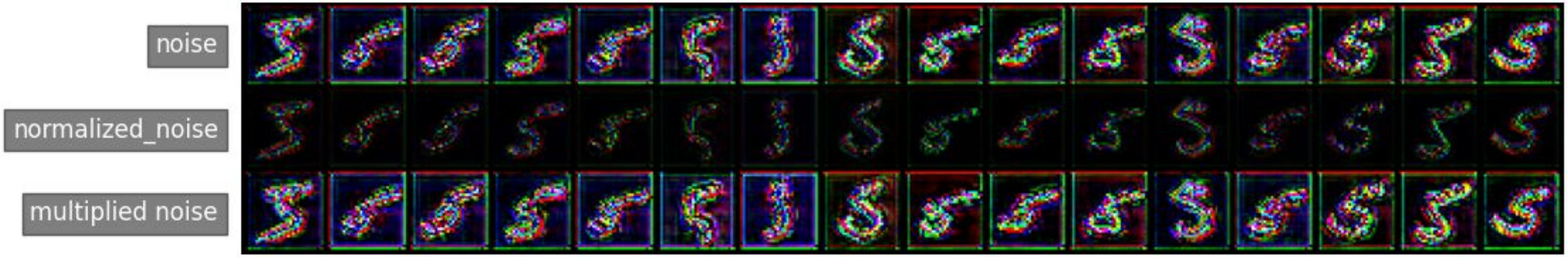


Sampling from normalized was bad (almost black result).Multiply by some coefficient:

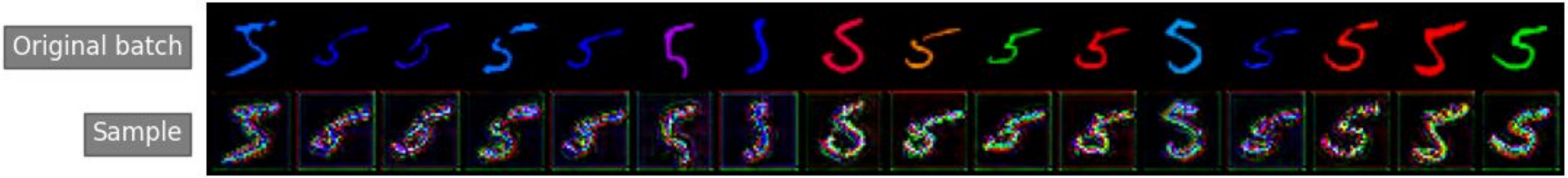


Experiments

Multiply by some coefficient:



And sample:



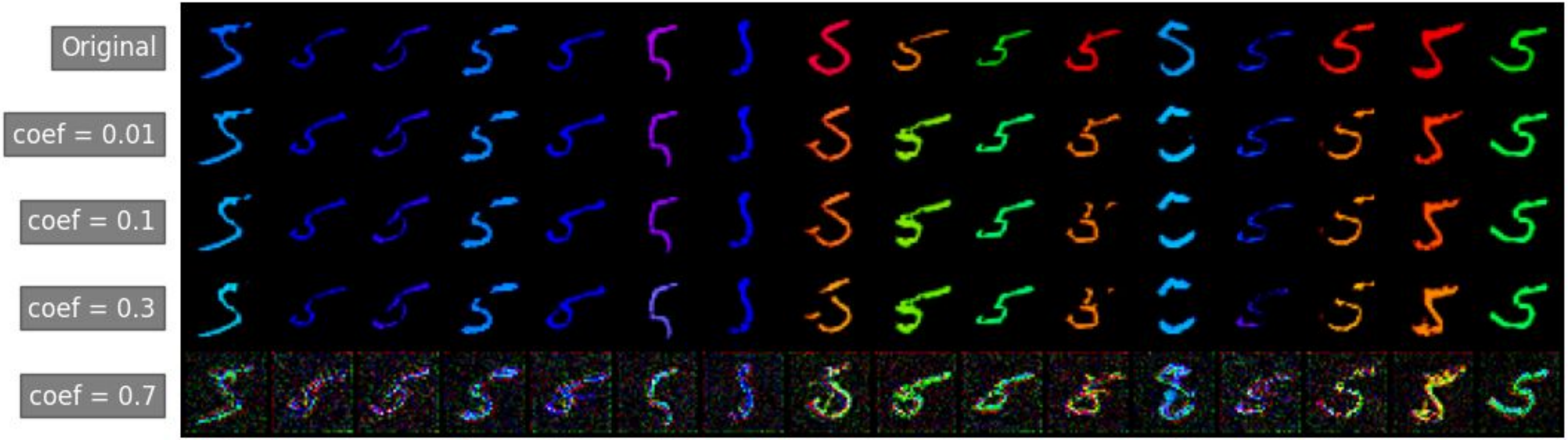
Experiments

1. Try to add some fixed Gauss noise to with some coefficient to:
   a. `noise=normalized_noise + c * fixed_noise` and iterate over `coefs = [0.01, 0.1, 0.3, 0.7]`

Experiments

`noise=noise + c * fixed_noise .` (source noise)

and iterate over `coefs = [0.01, 0.1, 0.3, 0.7]`



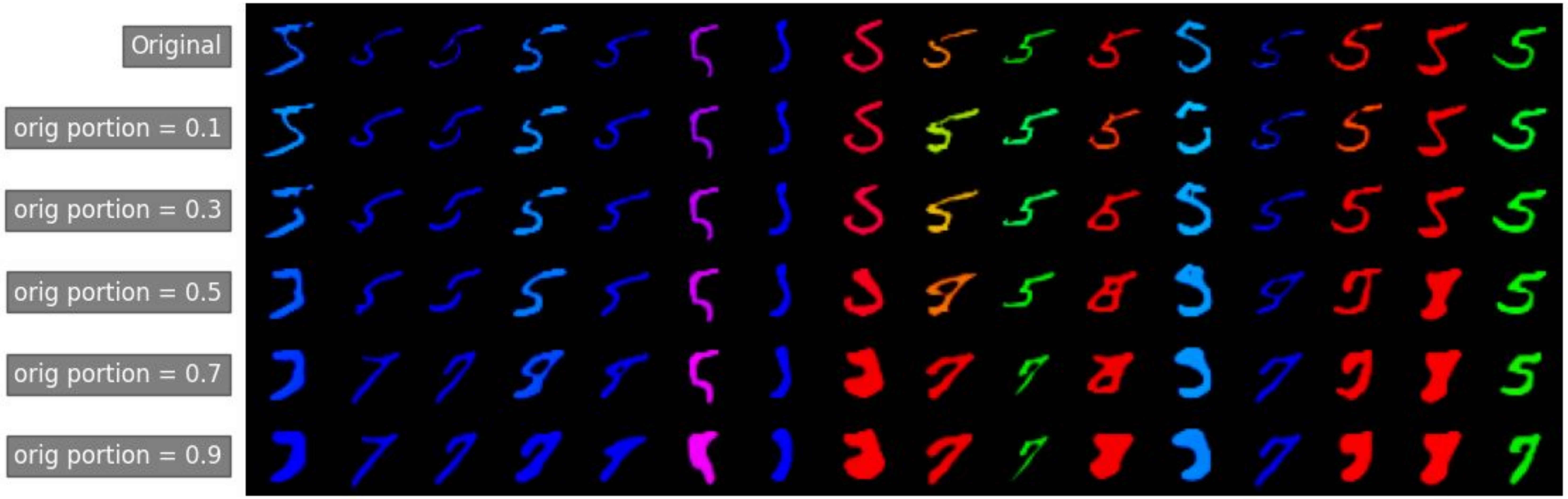
Experiments

Blending original image with its latent with different proportions:

```
proportions = [0.1, 0.3, 0.5, 0.7, 0.9]
```

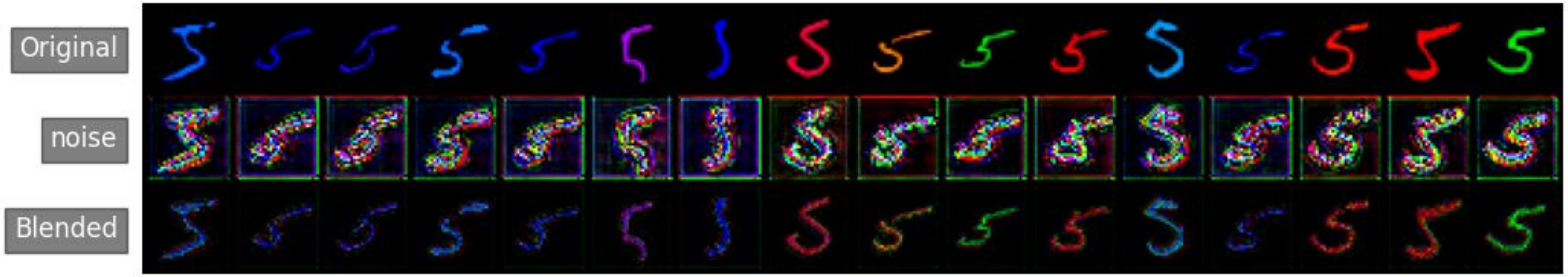
source noise:

Result:

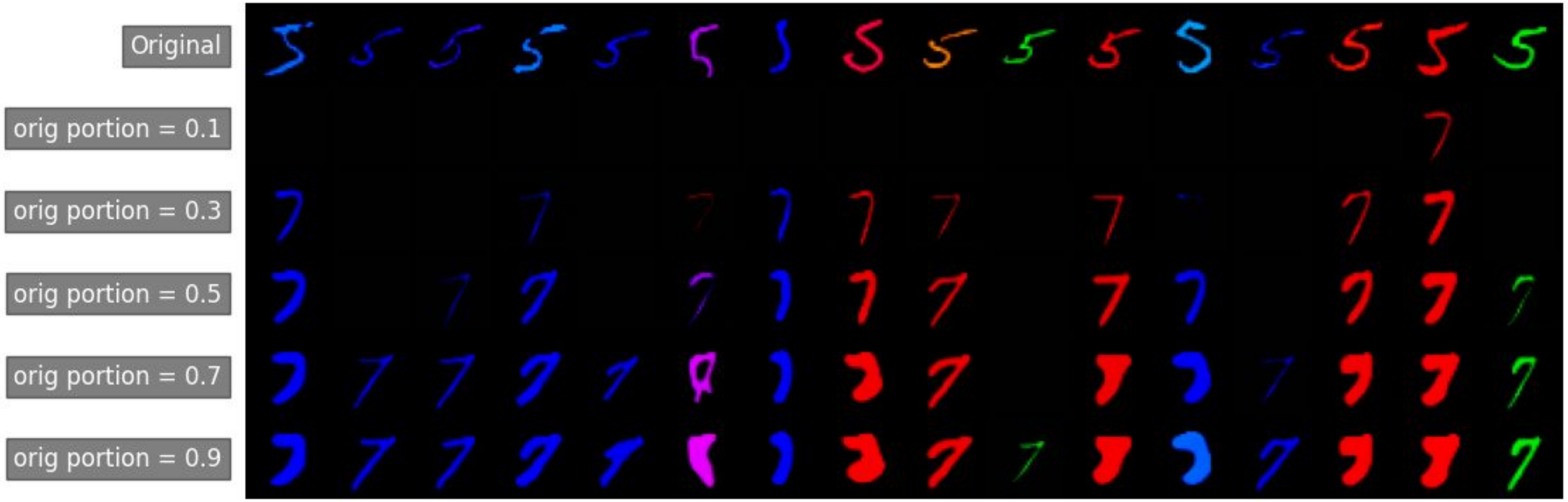


Experiments

normalized noise

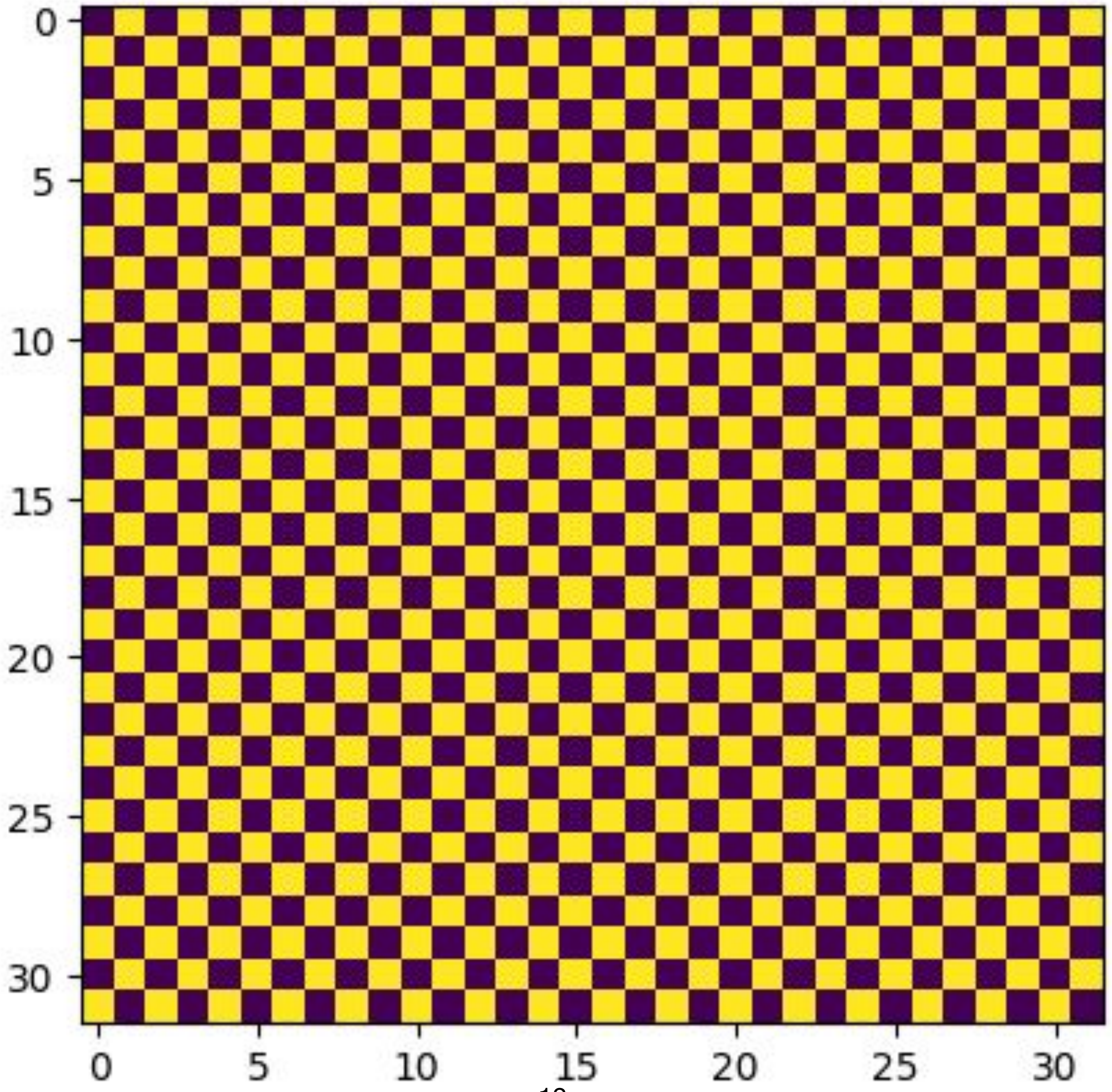


Result



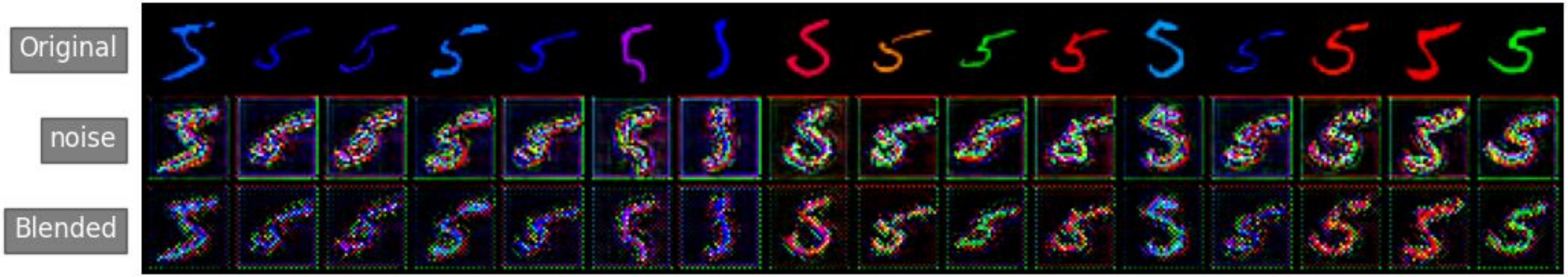
Experiments

Blending checkerboard style original image



Experiments

Source noise



Result

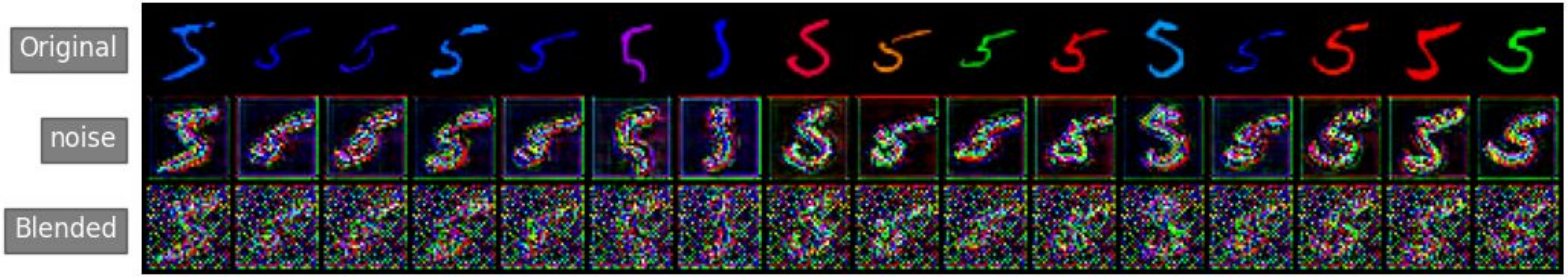
Experiments

Normalized noise

Result

Experiments

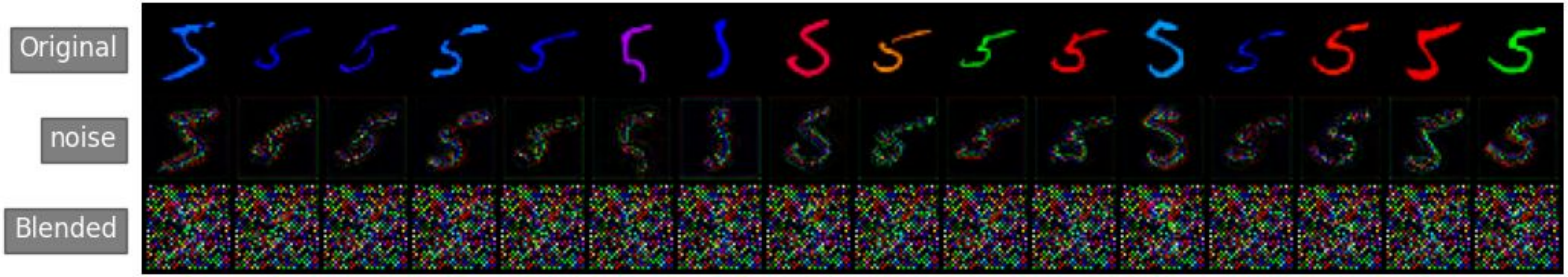
Blending in checkerboard style with fixed Gaussian nose
Source noise



Result

Experiments

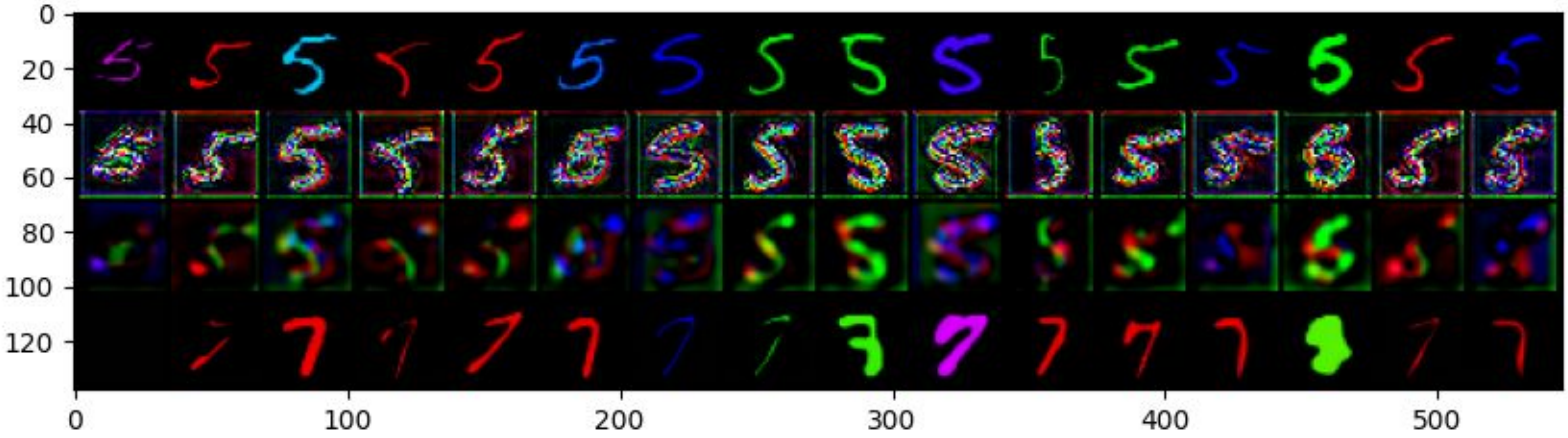
Normalized noise



Result

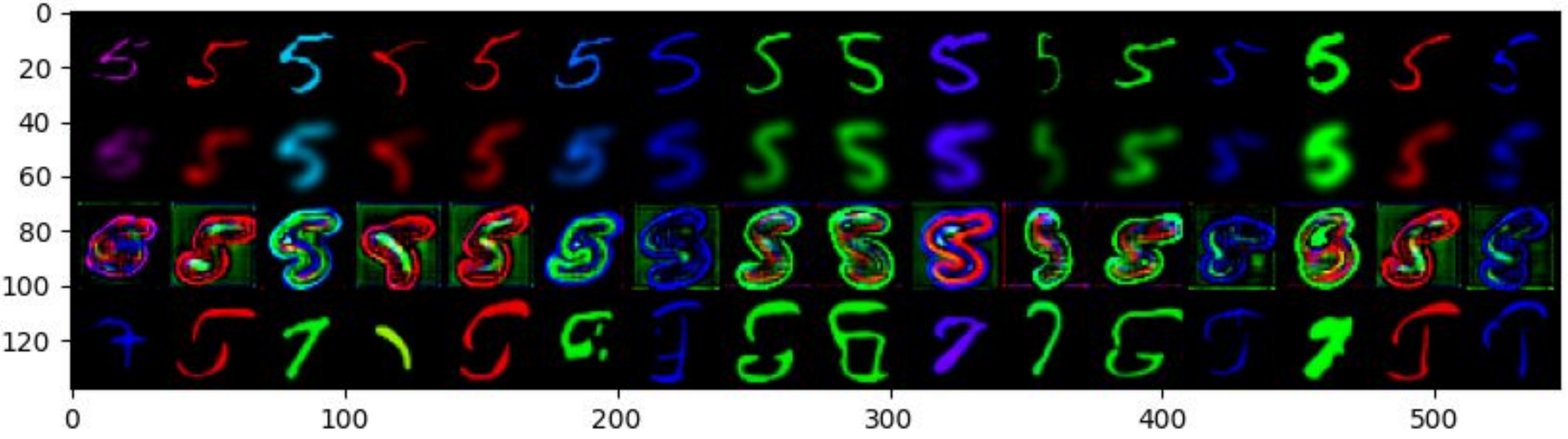
Additional experiments

First, the idea was to apply blur to the original image and see how it will go:



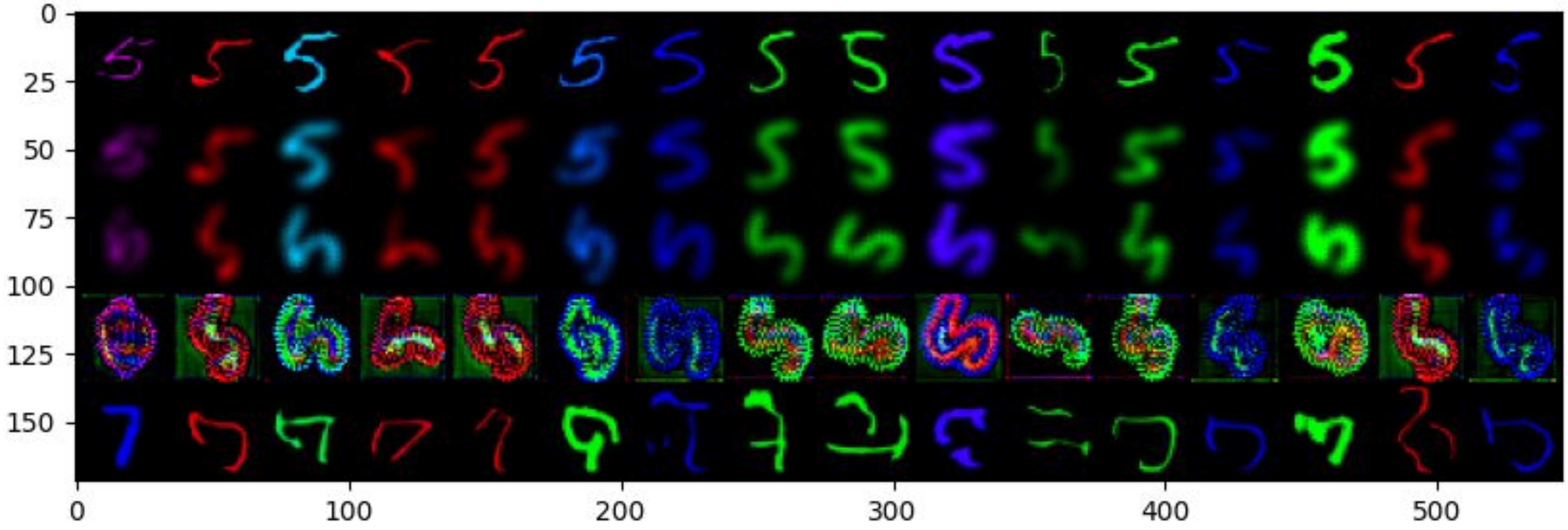
Additional experiments

Blur almost removed the structure of the original digit, however it also altered the color. Then we switched places and applied gaussian blur to the noise itself:



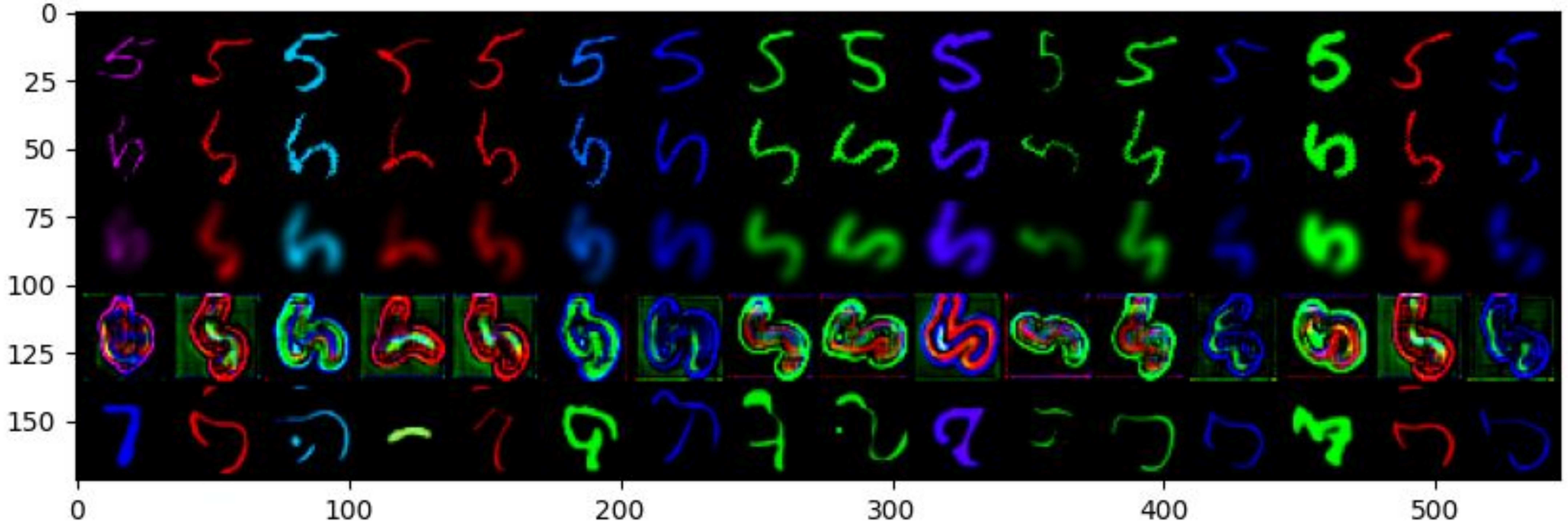
Additional experiments

Here we faced an interesting artefact: latent code gives a colorful contour to the digit. Now add rotation to the latent:



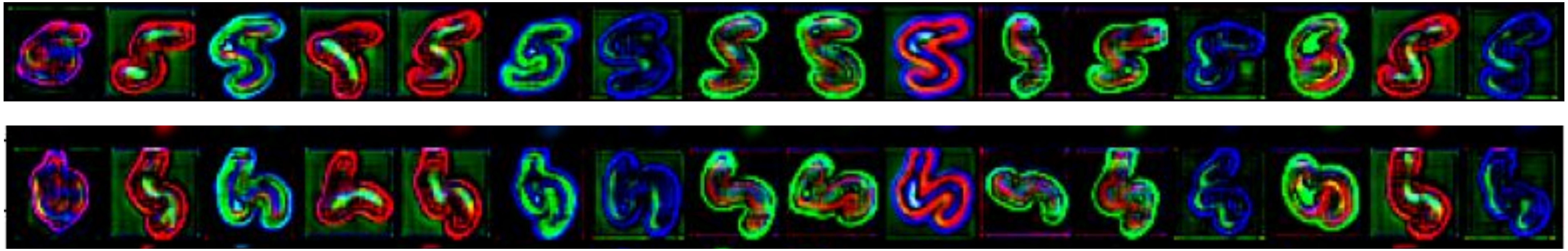
Additional experiments

Another interesting observation: the model is sensitive to the artefacts, which rotation creates. Now let's rotate and then blur:



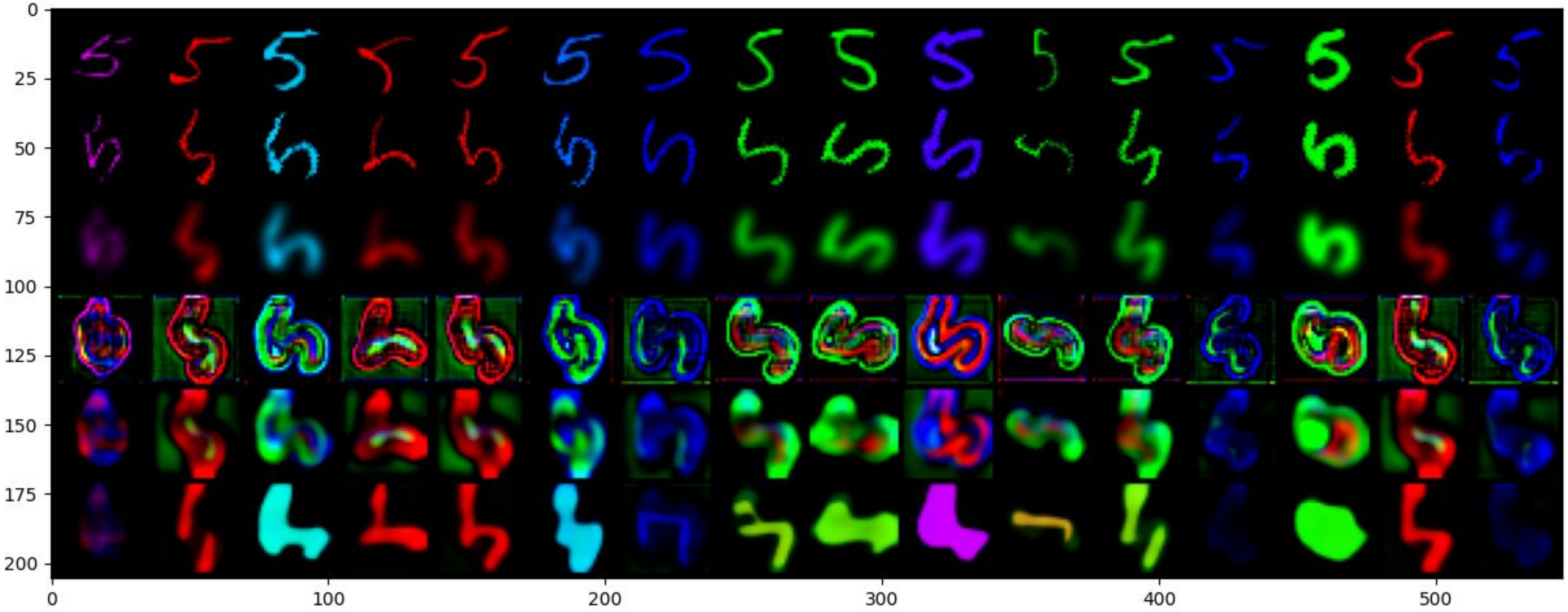
Additional experiments

The received latent is almost close to what we've got without rotation, thus we have slight equivariance



Additional experiments

After that we decided to blur the latent itself once more:



Well, we got the colors almost, but failed to get the needed structure of seven.

Conclusion

- Generated colored MNIST
- Trained a DDIB model to generate sevens
- Conducted experiments by getting the latent code of one digit and attempting to convert into seven, while preserving the colour
- Observed interesting results
- Github: https://github.com/MarioAuditore/image_to_image_ddpm

Team

Alexander Sharshavin

Elfat Sabitov

Regerences

1. Xuan Su, Jiaming Song, Chenlin Meng, Stefano Ermon, "Dual Diffusion Implicit Bridges for Image-to-Image Translation", 2023, ICLR'23
2. Nikita Gushchin, Alexander Kolesov, Alexander Korotin, Dmitry P. Vetrov, Evgeny Burnaev, "Entropic Neural Optimal Transport via Diffusion Processes", NeurIPS'23