

Porównanie modeli poprawy jakości obrazu (SR)

Kacper Jabłoński, Tymoteusz Dudek

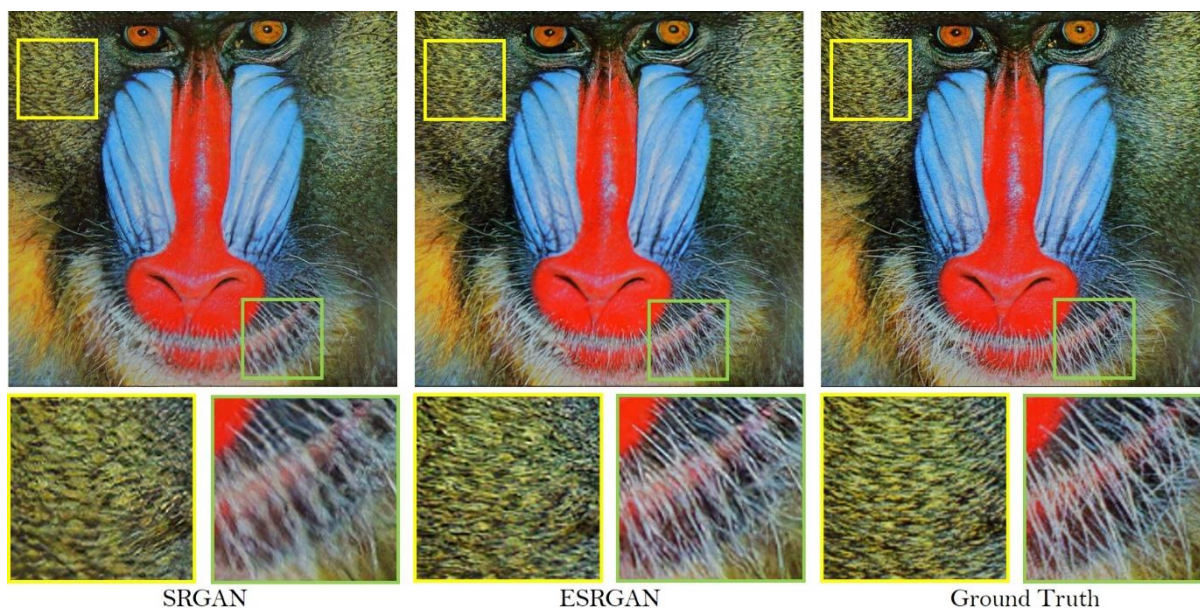
- [MemNet+](#)
- [ESRGAN](#)
- [Nasz github](#)
- [Przetworzone zdjęcia](#)

MemNet: A Persistent Memory Network for Image Restoration

Algorytm opiera się na głębokiej trwałej sieci pamięciowej (deep persistent memory network - MemNet), która przedstawia blok pamięci (*memory block*), składający się z jednostek: rekursywnej i bramowej (*gate*), w celu szerokiego wydobywania pamięci trwałej, poprzez adaptacyjny proces nauczania. Jednostka rekursywna uczy wielopoziomowej reprezentacji aktualnego stanu pod różnymi polami recepcyjnymi, natomiast reprezentacje oraz dane wyjściowe poprzednich bloków pamięci są scalane i wysyłane do jednostki bramowej, która adaptacyjnie kontroluje jak dużo poprzednich stanów powinno być zarezerwowanych i podejmuje decyzje jaka część aktualnego stanu powinna zostać zachowana. Stosujemy MemNet do trzech zadań, związanych z odzyskiwaniem obrazu: odszumianie obrazu, super rozdzielczości (super-resolution) oraz deblokoania JPEG.

ESRGAN

ESRGAN został stworzony jako udoskonalenie SRGAN'a. Poprawienie skupiało się na trzech aspektach. Pierwszym było wprowadzenie Residual-in-Residual Dense Block'ów oraz usunięcie warstw Batch Normalization, które były obciążeniem obliczeniowym i wpływały negatywnie na końcowy rezultat. Drugim była modyfikacja funkcjonalności sprawdzającej zbieżność obrazu z oczekiwaniami poprzez określanie realistyczności obrazu, a nie jego autentyczności. Trzecim elementem było wprowadzenie SR – MINC Loss, bazujące na VGG, które miało na celu lepiej oddawać stratę percepcyjną.



Porównanie modeli:

W tej części skupimy się na porównaniu algorytmów na tych samych danych. Algorytmy przetestujemy na zestawie danych, który nie był używany do ich trenowania. Skorzystamy ze zbioru danych, w wysokiej rozdzielczości, dla skali x4. Wykorzystamy 14 obrazów i porównamy je z obrazami uzyskanymi w wyniku działania modeli. Poniżej przedstawimy wyniki naszych testów dla poszczególnych algorytmów:

Zbiór testowy otrzymanych 14 obrazów

	MemNet	ESRGAN
Średni PSNR	32.26642278532077	31.606872599283292
Min PSNR	29.31840760102296	29.24473464408287
Max PSNR	35.36120880851898	33.89736000852671
Wariancja PSNR	3.263973938865391	2.2225892025611893
Odchylenie standardowe PSNR	1.806647153947165	1.4908350688661671
Średni SSIM	0.7124175747333517	0.6666920681683222
Min SSIM	0.3076302464285594	0.29885139517495246
Max SSIM	0.8969504490041827	0.8337591833914793
Wariancja SSIM	0.031036600482621284	0.02727166524814228
Odchylenie standardowe SSIM	0.1761720763419143	0.16514134929853966
Średni czas wykonania	3:44.883:0	11:42.85