



Wrocław University
of Science and Technology

Laboratorium 5

A. CHŁOPOWIEC, A. CHŁOPOWIEC, K. GALUS, W. CEBULA
254517, 254518, 253060, 256781

PROWADZĄCY: DR INŻ. PIOTR SZYMAŃSKI
KURS: PRZETWARZANIE DANYCH PRZESTRZENNYCH
GRUPA: CZWARTEK, 13:15

Spis treści

1	Opis zagadnienia	2
1.1	Sieci GAN w zadaniu Image-to-Image	2
1.2	Image-to-Image w Przetwarzaniu Danych Przestrzennych	2
1.3	Ekwiwariancja na grupę P4	3
2	Dane i eksperymenty	3
2.1	Dane	3
2.2	Eksperymenty	3
2.2.1	Opis eksperymentów	3
2.2.2	Wyniki	4
2.2.3	Wnioski - badania modeli	5
2.3	Wnioski - badania ekwiwariancji	6

1 Opis zagadnienia

1.1 Sieci GAN w zadaniu Image-to-Image

Zadanie przekształcenia Image-to-Image polega na przekształceniu próbki z domeny X na próbkę z rozkładu danych dziedziny Y . Taka transformacja może być wykonywana w dwóch ustawieniach: posiadamy pary $\{x, y\}$ i oczekujemy przekształcenia próbki x w próbkę y , lub posiadamy zbiór niepowiązanych próbek z dwóch domen. Wybraliśmy pierwsze podejście, jako że pozwala nam precyzyjnie zweryfikować jakość wykonywanych transformacji. Do tego celu zdecydowaliśmy się wykorzystać sieć GAN - Pix2Pix (1) jako podstawową warunkową sieć GAN, która w literaturze została wykorzystana do translacji danych przestrzennych.

1.2 Image-to-Image w Przetwarzaniu Danych Przestrzennych

W Przetwarzaniu Danych Przestrzennych istnieje potrzeba wykonywania przekształceń obrazu z jednej domeny do innej - w wielu przypadkach nie istnieją dobre mapy dla wielu miejsc na świecie, a w szczególności nie w różnych wariantach. Dokonywanie translacji w sposób automatyczny może zredukować koszty tworzenia takich map i podnieść ich ogólną jakość. Badania w tym zakresie są zazwyczaj przeprowadzane dla zdjęć o różnym poziomie przybliżenia.



Figure 1: Przykładowe obrazy dla dwóch domen z Google Maps, po lewej zdjęcia satelitarne, po prawej zdjęcia "roadmap" na przybliżeniach odpowiednio 18 i 16.

1.3 Ekwiwariancja na grupę $P4$

W naszym projekcie chcemy zbadać jakość modeli Pix2Pix opisanym wyżej zadaniu. Ponadto interesuje nas zbadanie ekwiwariancji na grupę $P4$ składającą się ze wszystkich kompozycji translacji i rotacji o 90 stopni. Z racji, że wykorzystujemy sieci konwolucyjne, mamy zapewnioną ekwiwariancję na translacje. Weryfikację ekwiwariancji na rotacje przeprowadziliśmy w trzech zestawieniach:

- Pix2Pix bez augmentacji w procesie trenowania.
- Pix2Pix z augmentacjami rotacji o $n \cdot 90$ stopni, gdzie $n=0,1,2,3$
- Pix2Pix ekwiwariantny na grupę $D4$ zaimplementowany z wykorzystaniem biblioteki `escnn` (2).

2 Dane i eksperymenty

2.1 Dane

Dane zebraliśmy przy użyciu API Google Maps Platform. Wybraliśmy dziedziny obrazów satelitarnych i typu "roadmap", w przybliżeniach 18 i 16. Dla każdej dziedziny i każdego przybliżenia zebraliśmy po 4000 obrazów, które podzieliliśmy na zbiór treningowy i walidacyjny w proporcjach 3:1. Wszystkie zdjęcia pochodzą z obszaru metropolitalnego Nowego Jorku, zdjęcia w zbiorze treningowym z części kontynentalnej, w zbiorze walidacyjnym z wyspy.

2.2 Eksperymenty

2.2.1 Opis eksperymentów

Wykonaliśmy eksperymenty dla każdego zestawienia:

- Model: Pix2Pix, Pix2Pix Rotation, Pix2Pix Equivariant
- Zadanie: Satellite - Roadmap, Roadmap - Satellite
- Przybliżenie: 16, 18

Każdy eksperyment dla modeli Pix2Pix i Pix2Pix Rotation został przeprowadzony na tym samym zestawie hiperparametrów. Jest to oficjalny zbiór parametrów wykorzystany w oryginalnej pracy.

Parametr	Wartość
Batch size	8
G_lr	0.0002
D_lr	0.0002
Image size	256
Lambda l1	100
Num epochs	500

Table 1: Tabela hiperparametrów dla modelu Pix2Pix, i Pix2Pix Rotate

Model ekwiwariantny został zaimplementowany w dwóch wariantach: z podstawową funkcją kosztu dla modelu Pix2Pix oraz z kosztem dla Wasserstein GAN z gradient penalty. Poniżej przedstawiamy zbiór wykorzystanych hiperparametrów.

Parametr	Wartość
Batch size	8
G_lr	0.0002
D_lr	0.0002
Image size	255
Lambda l1	100
Num epochs	500
Restriction	3
G - Channel div	1.6
D - Channel div	1.3
r1 subgroup	D2
r2 subgroup	D1

Table 2: Tabela hiperparametrów dla modelu Pix2Pix Equivariant

2.2.2 Wyniki

Modele Pix2Pix oraz Pix2Pix Rotate nauczyły się dosyć dobrze przekształcać dane z domeny X do Y, w obu zadaniach i na obu przybliżeniach. Nie byliśmy w stanie wyuczyć modelu ekwiwariantnego, zarówno w formie podstawowej jak, WGAN-GP. Prawdopodobnie jest to spowodowane błędem na poziomie architektury, jednakże nie udało nam się go zidentyfikować. Poniżej przedstawiamy wyniki dla modeli Pix2Pix oraz Pix2Pix Rotate.

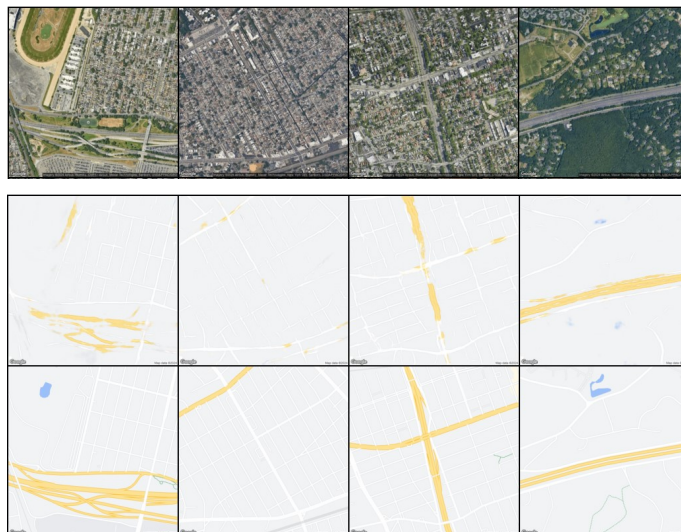


Figure 2: Pix2Pix Satellite - Roadmap Zoom 16. Pierwszy rząd, dane satelitarne, drugi rząd, dane wygenerowane, trzeci rząd, dane rzeczywiste "roadmap".

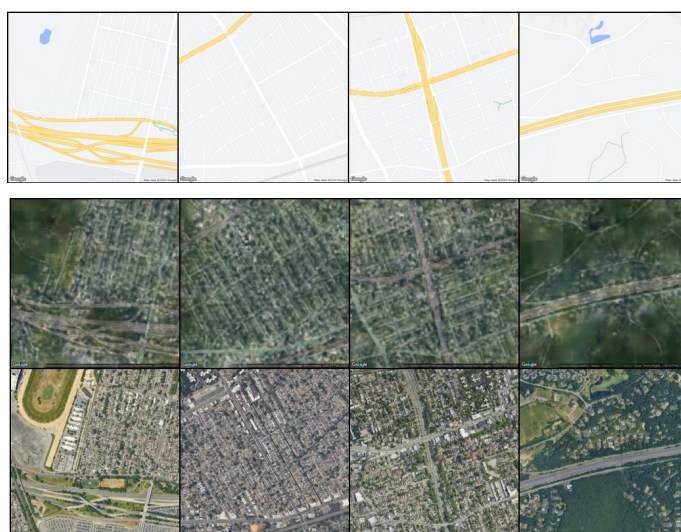


Figure 3: Pix2Pix Roadmap - Satellite Zoom 16. Pierwszy rząd, dane "roadmap", drugi rząd, dane wygenerowane, trzeci rząd, dane rzeczywiste satelitarne.

2.2.3 Wnioski - badania modeli

Modele typu GAN mają duży problem z nauczeniem się odtwarzania prostych linii. Ponadto często nie były w stanie rozróżnić z jakiego typu drogą mamy do czynienia - autostradą/drogą ekspresową, czy zwykłą drogą miejską. Kod jest dostępny na repozytorium github: <https://github.com/Adam-Chlopowiec/equivariant-map-gen>

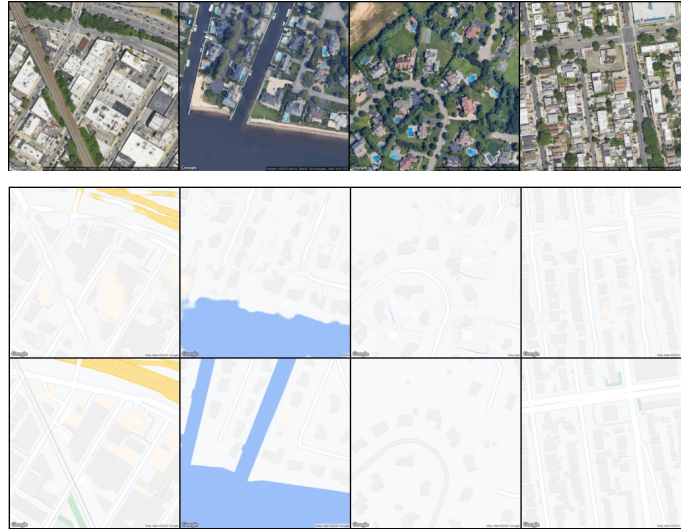


Figure 4: Pix2Pix Satellite - Roadmap Zoom 18. Pierwszy rząd, dane satelitarne, drugi rząd, dane wygenerowane, trzeci rząd, dane rzeczywiste "roadmap".

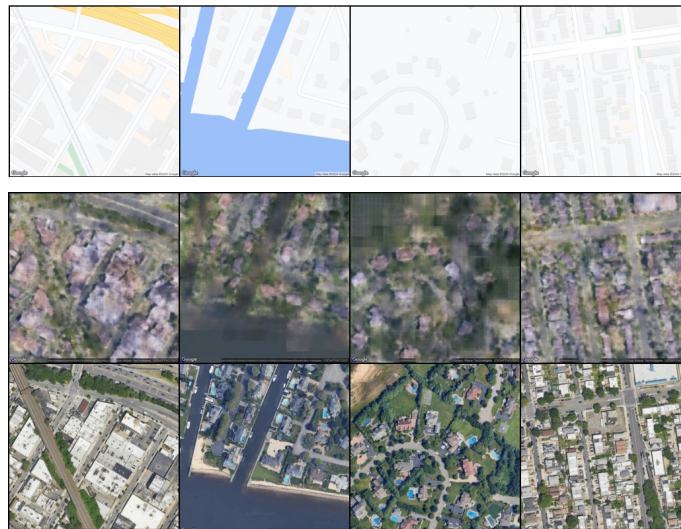


Figure 5: Pix2Pix Roadmap - Satellite Zoom 18. Pierwszy rząd, dane "roadmap", drugi rząd, dane wygenerowane, trzeci rząd, dane rzeczywiste satelitarne.

2.3 Wnioski - badania ekwiwariancji

Wśród osób zaczynających zajmowanie się zagadnieniami symetrii i grupowej ekwiwariancji często pojawia się dyskurs dotyczący celowości nakładania ograniczeń na przestrzeń wyuczalnych funkcji względem zastosowania augmentacji danych. W literaturze można znaleźć informacje jakoby wprowadzenie augmentacji rotacji w proces

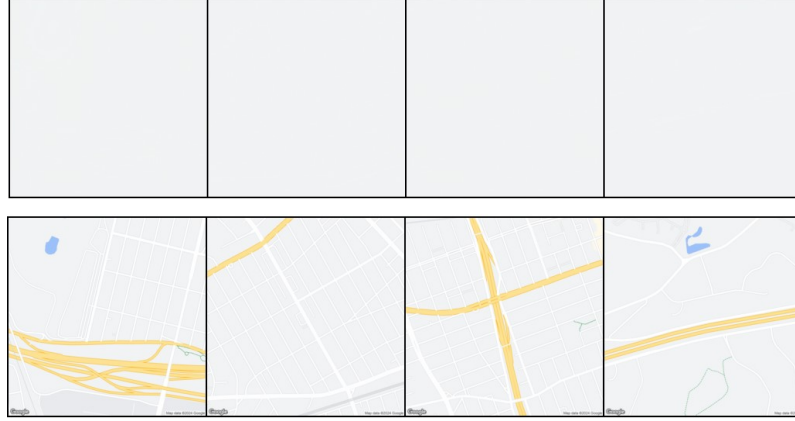


Figure 6: Wyniki źle wyuczonego model Pix2Pix Equivariant

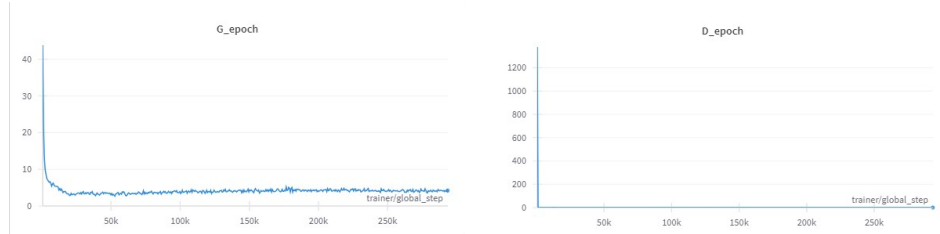


Figure 7: Krzywe uczenia źle wyuczonego model Pix2Pix Equivariant

uczenia miało uczyć sieć ekwiwariancji na tą grupę. Głównym powodem stosowania grupowo ekwiwariantnych sieci neuronowych jest gwarancja dotycząca ekwiwariancji, która umożliwia stabilną generalizację w tym zakresie. Podobne wnioski jak te wyciągnięte z powyższego eksperymentu znajdują również odzwierciedlenie w literaturze. Przygotowaliśmy zbiór obrazów, na którym zastosowaliśmy obrót o 90, 180 i 270 stopni. Każdy z tych obrazów został przetworzony przez wyuczoną sieć, a następnie obrócony z powrotem do orientacji kanonicznej. Dla każdego obrotu policzyliśmy MSE między obrazem w orientacji kanonicznej przetworzonym przez sieć, a obrazami zrotowanymi przetworzonymi przez sieć, obróconymi do postaci kanonicznej. Ponadto zebraliśmy wyniki wizualizacji translacji. Z przeprowadzonych badań zaobserwowaliśmy, że augmentacja danych w procesie uczenia jest niewystarczająca do wprowadzenia ekwiwariancji na rotacje do modelu. Niemniej jednak model uczony z augmentacjami osiąga mniejsze błędy, co pokazuje, że nie jest to zabieg bez celowy.

References

- [1] Isola, Philip. *Image-to-Image Translation with Conditional Adversarial Networks*. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017.
- [2] Gabriele Cesa. *A Program to Build $E(N)$ -Equivariant Steerable CNNs*. International Conference on Learning Representations, 2022

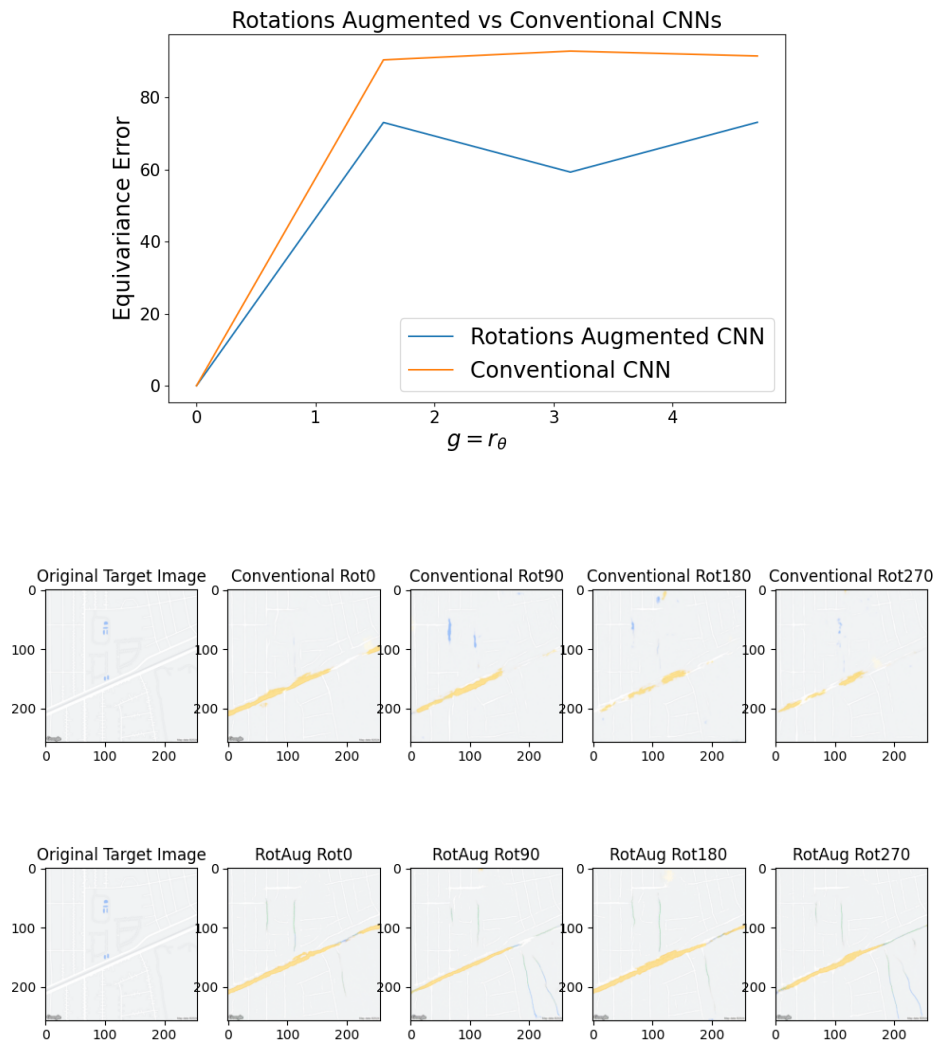


Figure 8: Pix2Pix Satellite - Roadmap Zoom 16. Wykres błędu i porównanie rekonstrukcji dla modelu. W pierwszym wierszu model bez augmentacji z rotacjami, w drugim wierszu model uczony z augmentacjami.

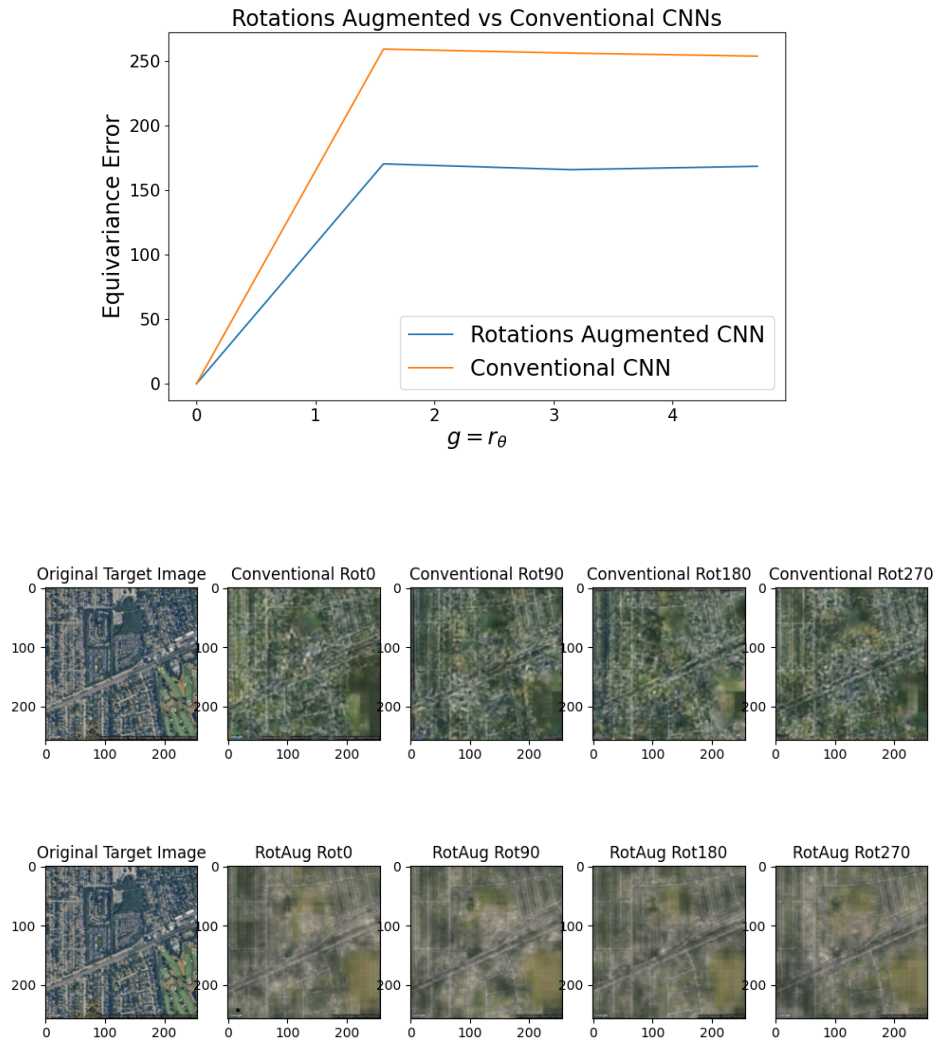


Figure 9: Pix2Pix Roadmap - Satellite Zoom 16. Wykres błędu i porównanie rekonstrukcji dla modelu. W pierwszym wierszu model bez augmentacji z rotacjami, w drugim wierszu model uczony z augmentacjami.

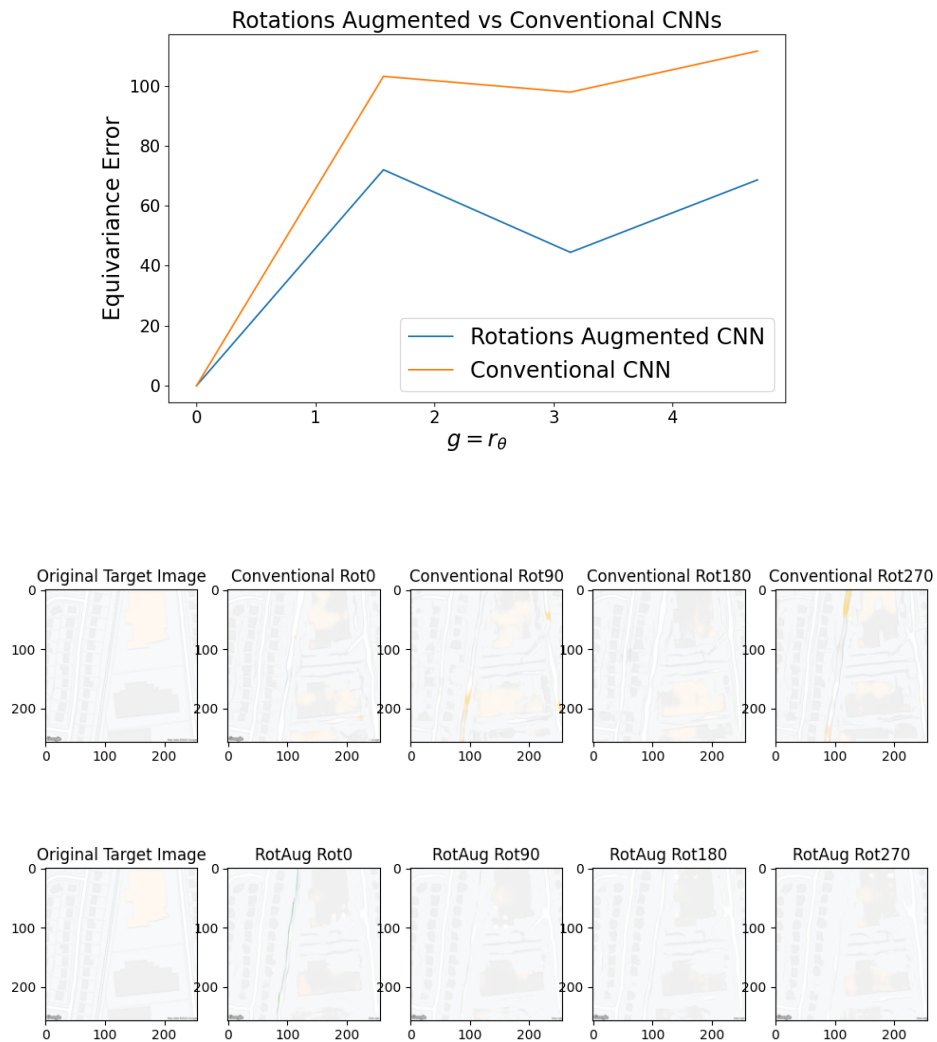


Figure 10: Pix2Pix Satellite - Roadmap Zoom 18. Wykres błędu i porównanie rekonstrukcji dla modelu. W pierwszym wierszu model bez augmentacji z rotacjami, w drugim wierszu model uczony z augmentacjami.

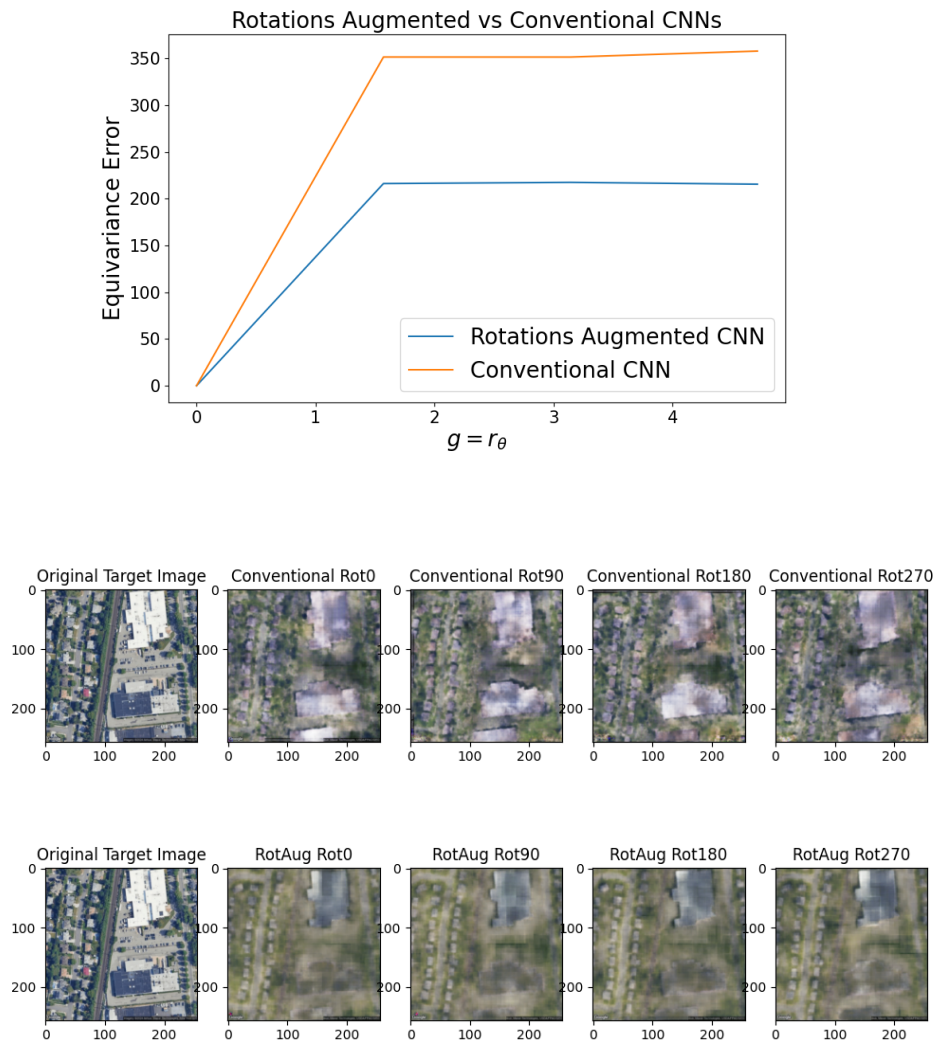


Figure 11: Pix2Pix Roadmap - Satellite Zoom 18. Wykres błędu i porównanie rekonstrukcji dla modelu. W pierwszym wierszu model bez augmentacji z rotacjami, w drugim wierszu model uczony z augmentacjami.