



# Wzbogacanie danych w analizie obrazów

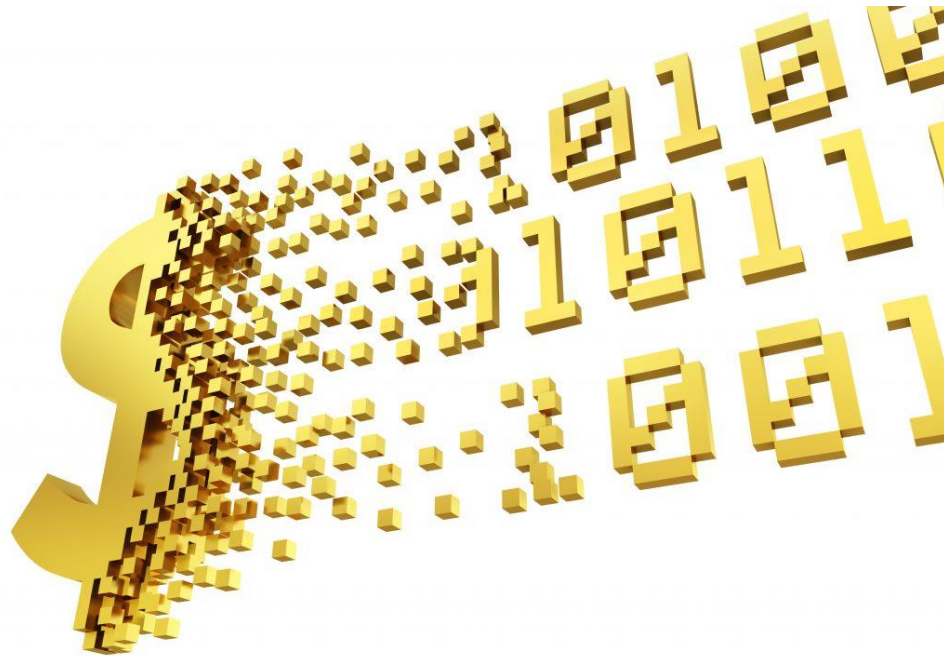
Stanisław Kaźmierczak



# Agenda

- Wprowadzenie
- Taksonomia
- Przekształcenia geometryczne
- Przekształcenia koloru
- Metody zaawansowane
- Wyniki, wskazania
- Podsumowanie

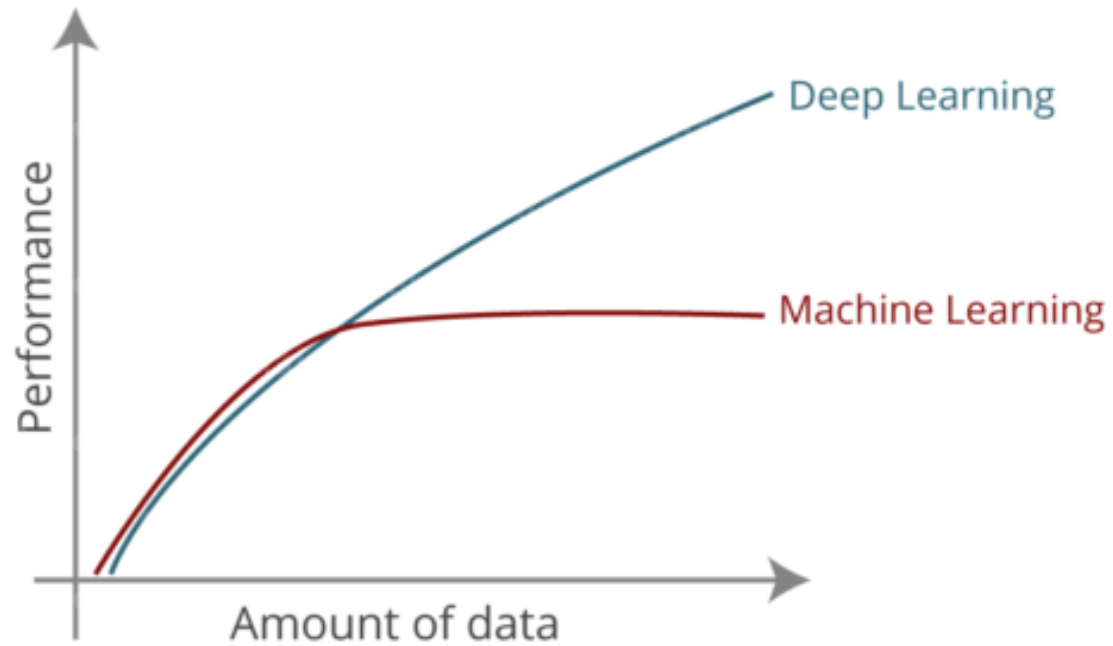
# Wprowadzenie



[1]

*Data augmentation in computer vision*

# Wprowadzenie



[2]



# Wprowadzenie

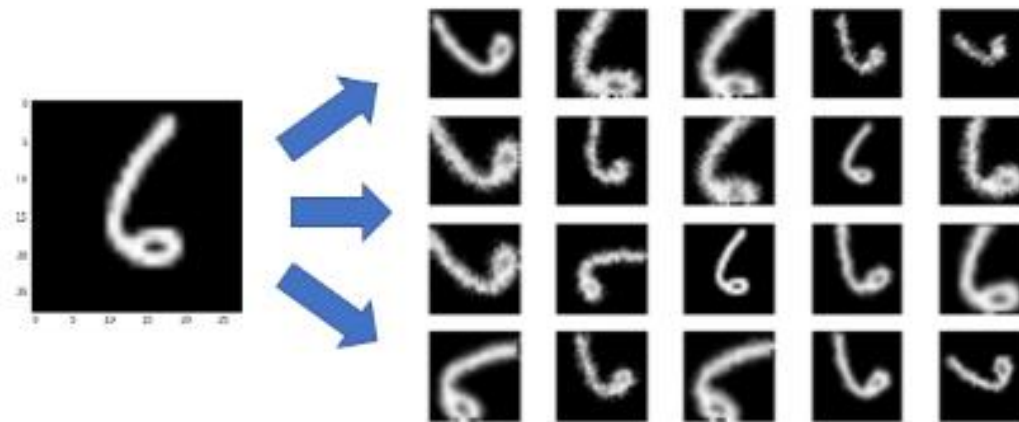
- Etykietyzacja danych stanowi przeważnie problem
  - Jest kosztowna
  - Jest czasochłonna
  - Wymaga znacznego zaangażowania ekspertów dziedzinowych
- Sieci neuronowe osiągające najlepsze rezultaty w obszarze analizy obrazów posiadają parametry, których liczność jest rzędu milionów, dziesiątek milionów, a nawet setek milionów
- Wymusza to, aby zbiór treningowy był bardzo liczny i możliwie zróżnicowany
  - Ręczna etykietyzacja zbiorów tej wielkości jest niemal niemożliwa



# Wprowadzenie

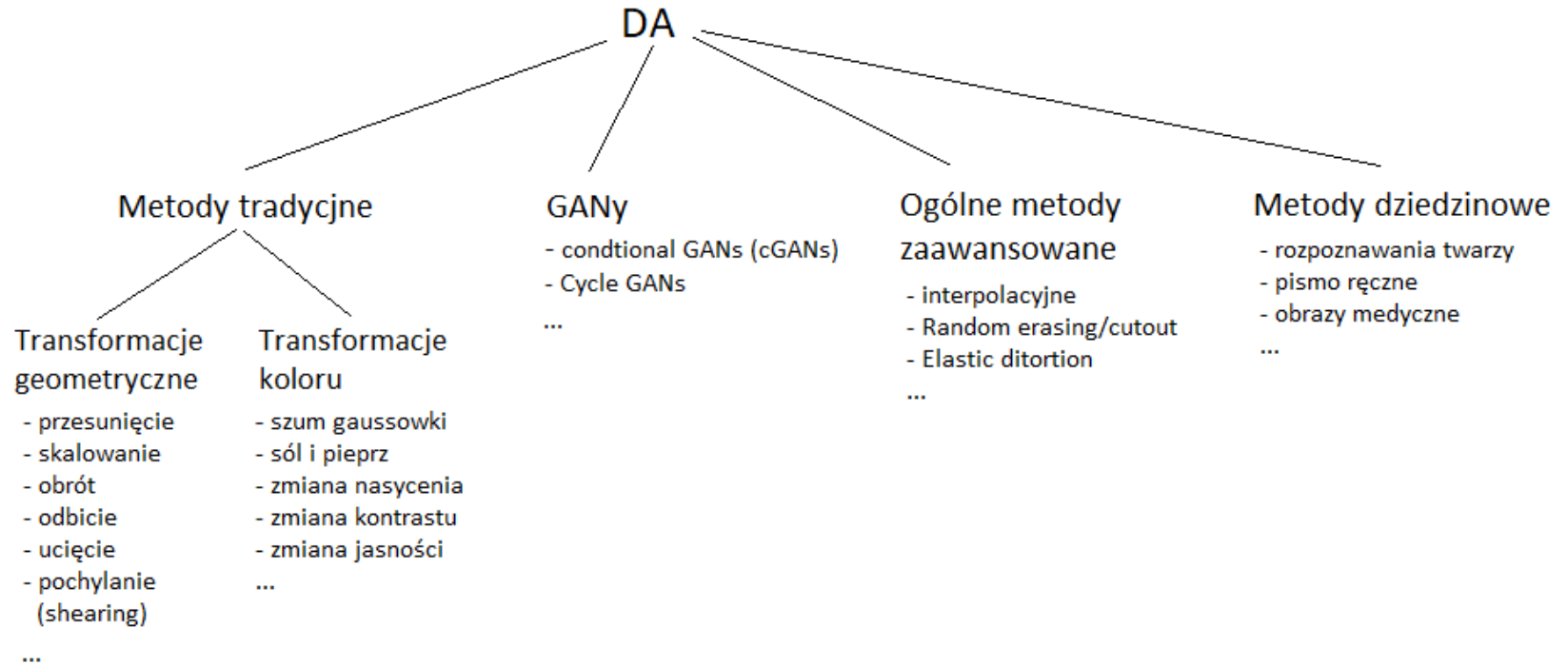
- Ogólna idea: przekształcić obraz w ten sposób, aby był on istotnie różny od oryginału, a jednocześnie wciąż należał do tej samej klasy
- Augmentacja danych (DA) może uczynić sieć odporną względem bazowej orientacji obiektów
  - Warunki, w których działa sieć mogą być dużo bardziej zróżnicowane niż warunki, w których tworzony był zbiór treningowy (inne oświetlenie, inny kąt patrzenia, itp.)
- Augmentacja pomaga zapobiec sytuacji, w której sieć uczy się niewłaściwych wzorców
  - Przykład: samochodom marki Ford towarzyszy słoneczna pogoda, a samochodom marki Opel – deszczowa

# Wprowadzenie



[3]

# Taksonomia







# Rodzaje DA

- Offline DA
  - Wszelkie przekształcenia robione są przed rozpoczęciem treningu sieci
  - Preferowany dla małych zbiorów małych rozmiarów
- Online DA
  - Transformacje robione są na bieżąco, bezpośrednio przed podaniem próbek do sieci
  - Wydajniejsze pamięciowo
  - Nie każdą metodę DA da się przeprowadzić w tym trybie, np. techniki, w których syntetyczny obraz powstał na bazie wielu obrazów

# Macierz transformacji

$$\begin{bmatrix} 1 & 0 & T_x \\ 0 & 1 & T_y \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & T_x \\ 0 & 1 & T_y \end{bmatrix}$$

**A**   **B**

$$\begin{bmatrix} x_{new} \\ y_{new} \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} + B$$

$$\begin{bmatrix} x_{new} \\ y_{new} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix}$$

[7]

# Flipping

- Pionowy lub poziomy
- Współczynnik DA (*DA Factor*): 4



[3]

$$\left[ \begin{array}{cc|c} 1 & 0 & 0 \\ 0 & -1 & 0 \end{array} \right]$$

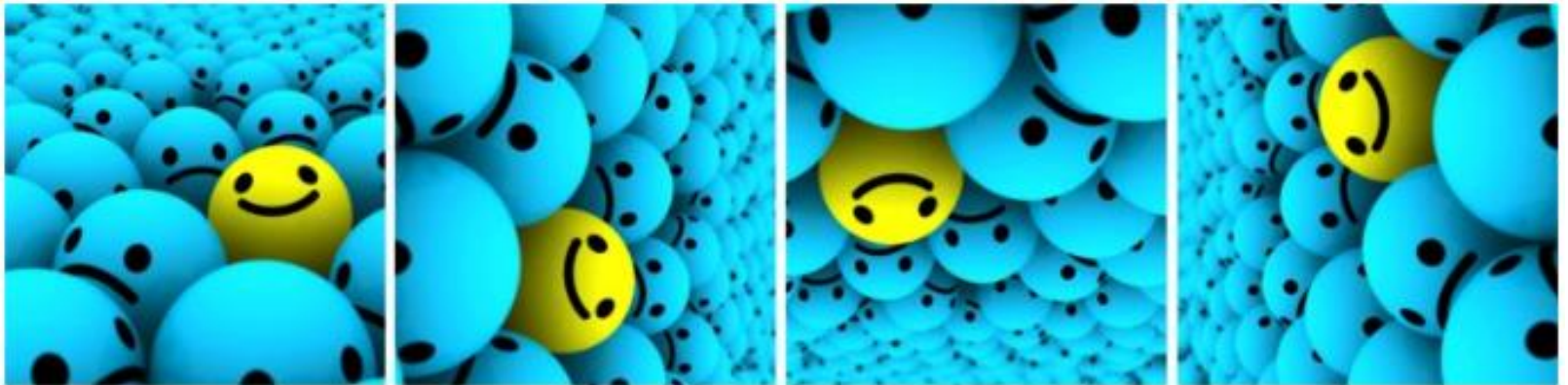
Odbicie wzdłuż osi x

$$\left[ \begin{array}{cc|c} -1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right]$$

Odbicie wzdłuż osi y

# Obrót

- Poza kilkoma szczególnymi kątami obrotu, rozmiar obrazka ulega zmianie
- Współczynnik DA: nieograniczony



[3]

$$\left[ \begin{array}{cc|c} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \end{array} \right]$$

Macierz obrotu

# Skalowanie

- Polega na zmianie rozmiaru obrazu
- Z reguły przetworzony obraz powinien być tego samego rozmiaru co oryginalny
  - Zwiększanie rozmiaru powoduje konieczność ucięcia obrazu
  - Zmniejszanie wymusza podjęcia dodatkowych działań związanych z powstałą ramką
- Współczynnik DA: nieograniczony



[3]

$$\left[ \begin{array}{cc|c} S_x & 0 & 0 \\ 0 & S_y & 0 \end{array} \right]$$

Macierz skalowania

# Cropping

- Często towarzyszy innym przekształceniom
- Istnieje również jako samodzielna metoda – *random cropping*
- Współczynnik DA: nieograniczony



[3]

# Przesunięcie

- Stymuluje sieć do „szukania” obiektu na całym obrazie
- Współczynnik DA: nieograniczony



[3]

$$\left[ \begin{array}{cc|c} 1 & 0 & T_x \\ 0 & 1 & T_y \end{array} \right]$$

Macierz przesunięcia

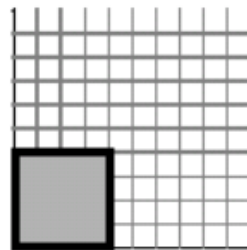
# Shearing

- Liniowe mapowanie, w którym:
  - wszystkie punkty należące do ustalonej prostej  $L$  nie zmieniają swojego położenia
  - wszystkie pozostałe punkty są przesuwane w kierunku równoległym do  $L$  o wielkość proporcjonalną do odległości punktu od  $L$
- Współczynnik DA: nieograniczony

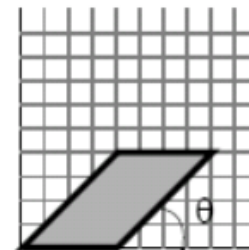
$$\left[ \begin{array}{cc|c} 1 & Sh_x & 0 \\ Sh_y & 1 & 0 \end{array} \right]$$

Macierz pochylenia

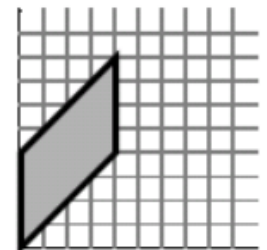
original



x - shear



y - shear



[8]

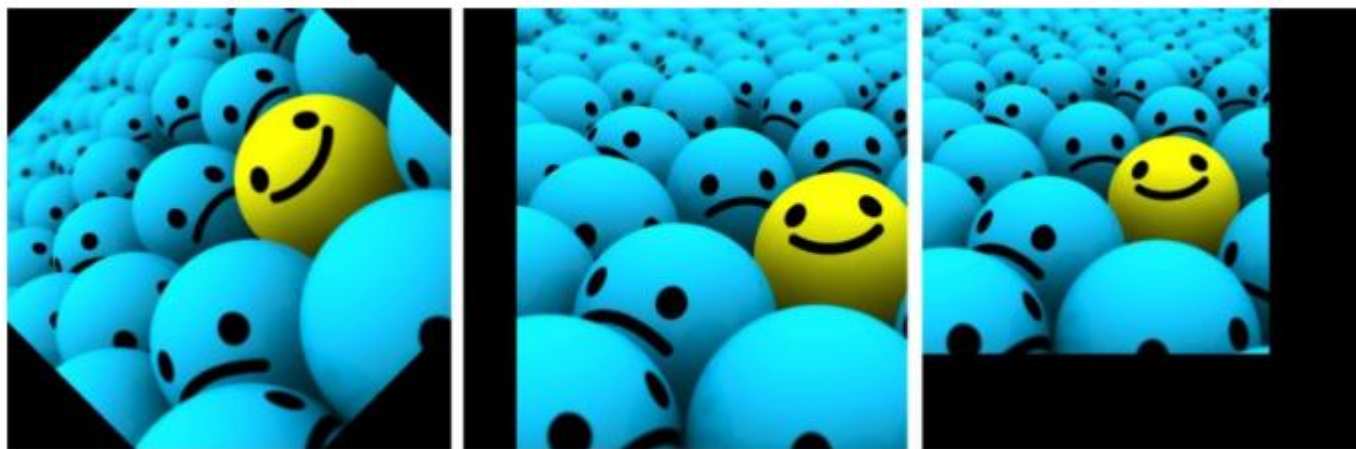


# Przekształcenia geometryczne

Operacja	Macierz przekształcenia
Obrót	$\left[ \begin{array}{cc c} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \end{array} \right]$
Odbicie	$\left[ \begin{array}{cc c} 1 & 0 & 0 \\ 0 & -1 & 0 \end{array} \right] \quad \left[ \begin{array}{cc c} -1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right]$
Skalowanie	$\left[ \begin{array}{cc c} S_x & 0 & 0 \\ 0 & S_y & 0 \end{array} \right]$
Przesunięcie	$\left[ \begin{array}{cc c} 1 & 0 & T_x \\ 0 & 1 & T_y \end{array} \right]$
Pochylenie	$\left[ \begin{array}{cc c} 1 & Sh_x & 0 \\ Sh_y & 1 & 0 \end{array} \right]$

# Problem pustych powierzchni

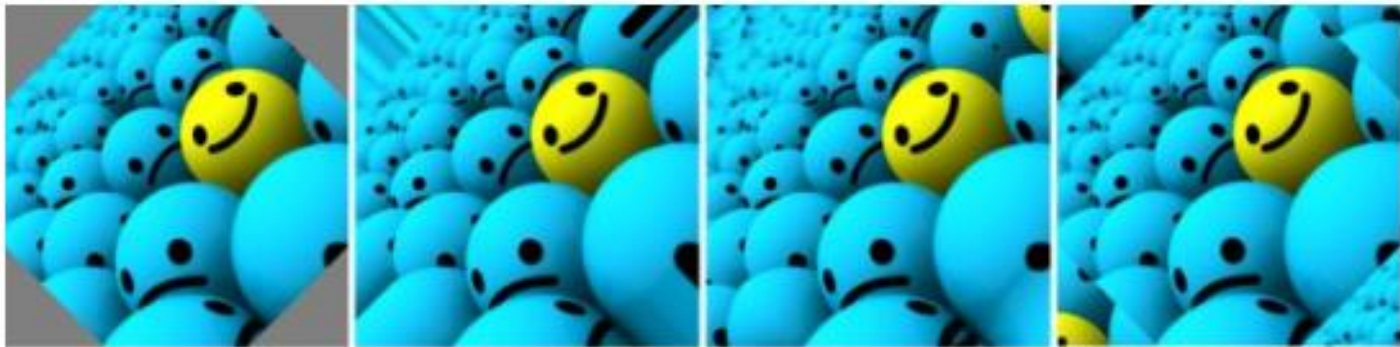
- W wyniku większości omawianych przekształceń powstają puste miejsca
- Najprostszym sposobem poradzenia sobie z tym problemem jest ucięcie, następnie przeskalowanie obrazu
- Bardziej zaawansowanym - interpolacja



[3]

# Interpolacja

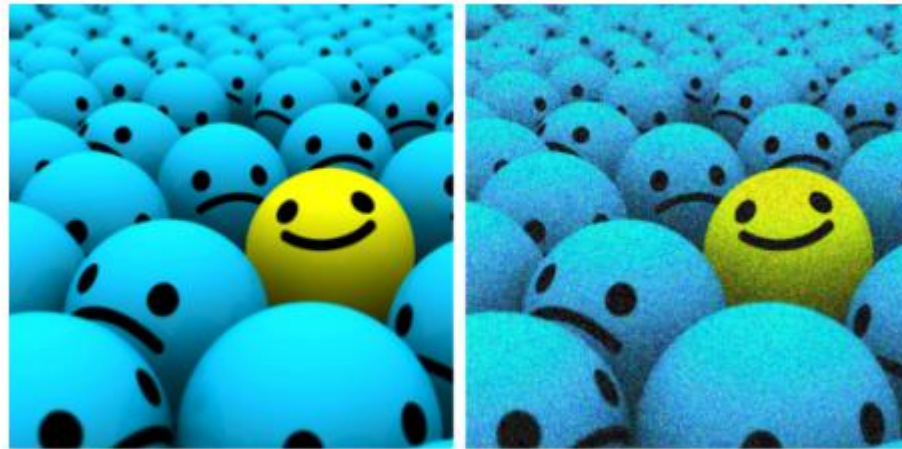
- Stała – puste miejsce jest zastępowane stałym kolorem (szczególnie użyteczna dla obrazów monochromatycznych)
- Brzegowa – puste miejsce jest zastępowane wartościami brzegowymi
- Odbicie – puste miejsce jest zastępowane wartościami pikseli odbitych wzdłuż krawędzi
- Zawijanie – puste miejsce jest zastępowane częścią obrazu z przeciwnej strony



[3]

# Szum gaussowski

- Zmiana jasności pikseli o wartość zmiennej losowej z rozkładu normalnego



[3]

# Sól i pieprz

- Zmiana koloru losowo wybranych pikseli na biały lub czarny

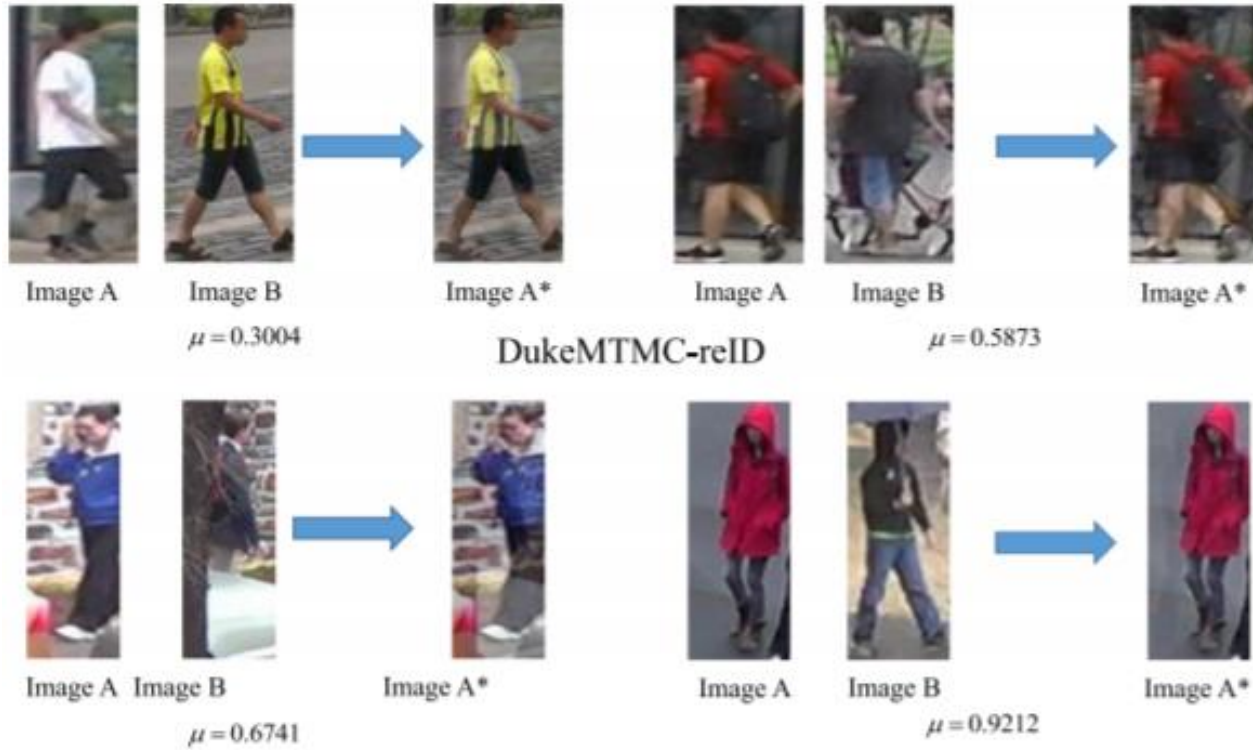


[4]



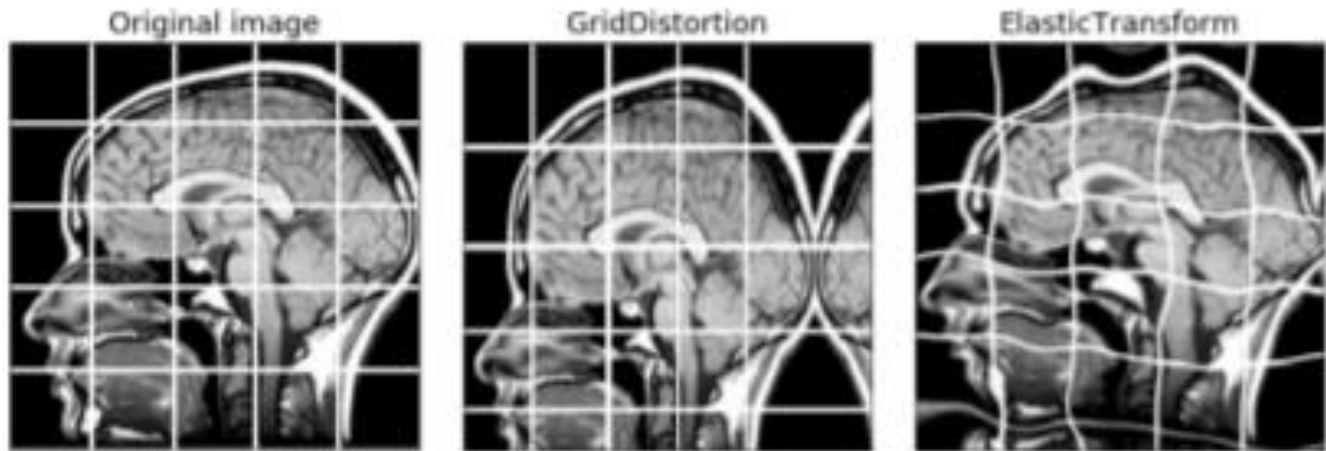


# Interpolacja liniowa



[11]

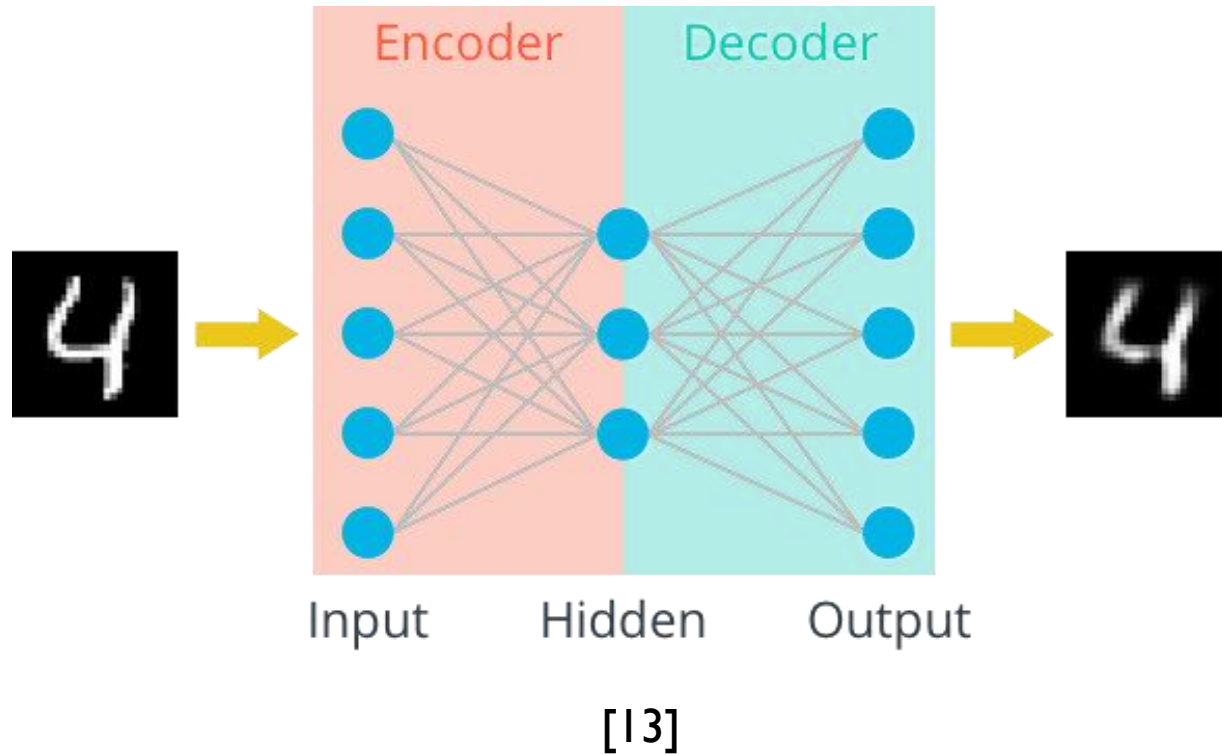
# Deformacije



[12]



# Autoenkodery



# Wskazówki

- DA powinno rozszerzać zbiór tylko o realne przypadki



Oryginał



Flip poziomy, przydatny dla zadania klasyfikacji samochodów na drodze



Obrót o 180 st. ma zastosowanie w przypadku zadań związanych z wypadkami drogowymi



Obrót o 90 st. nie ma zastosowania do żadnego z wymienionych zadań



# Wskazówki

- Należy zadbać, aby metody DA nie zmieniały klasy obiektu (np. obrót o 180 st. cyfry „9” lub flip litery „b”)
- DA w przestrzeni cech
  - Ogólne, niezależne od dziedziny
- DA w przestrzeni danych
  - Związane z konkretną dziedziną
- Rozsądna (zachowująca etykiety) augmentacja w przestrzeni danych przynosi lepsze rezultaty [6]

# Wyniki dla CIFAR10 i CIFAR100

Method	Depth	Params	C10	C10+	C100	C100+
Network in Network [22]	-	-	10.41	8.81	35.68	-
All-CNN [32]	-	-	9.08	7.25	-	33.71
Deeply Supervised Net [20]	-	-	9.69	7.97	-	34.57
Highway Network [34]	-	-	-	7.72	-	32.39
FractalNet [17]	21	38.6M	10.18	5.22	35.34	23.30
with Dropout/Drop-path	21	38.6M	7.33	4.60	28.20	23.73
ResNet [11]	110	1.7M	-	6.61	-	-
ResNet (reported by [13])	110	1.7M	13.63	6.41	44.74	27.22
ResNet with Stochastic Depth [13]	110	1.7M	11.66	5.23	37.80	24.58
	1202	10.2M	-	4.91	-	-
Wide ResNet [42]	16	11.0M	-	4.81	-	22.07
	28	36.5M	-	4.17	-	20.50
with Dropout	16	2.7M	-	-	-	-
ResNet (pre-activation) [12]	164	1.7M	11.26*	5.46	35.58*	24.33
	1001	10.2M	10.56*	4.62	33.47*	22.71
DenseNet ( $k = 12$ )	40	1.0M	<b>7.00</b>	5.24	<b>27.55</b>	24.42
DenseNet ( $k = 12$ )	100	7.0M	<b>5.77</b>	<b>4.10</b>	<b>23.79</b>	<b>20.20</b>
DenseNet ( $k = 24$ )	100	27.2M	<b>5.83</b>	<b>3.74</b>	<b>23.42</b>	<b>19.25</b>
DenseNet-BC ( $k = 12$ )	100	0.8M	<b>5.92</b>	4.51	<b>24.15</b>	22.27
DenseNet-BC ( $k = 24$ )	250	15.3M	<b>5.19</b>	<b>3.62</b>	<b>19.64</b>	<b>17.60</b>
DenseNet-BC ( $k = 40$ )	190	25.6M	-	<b>3.46</b>	-	<b>17.18</b>

[5]



# Podsumowanie

- DA poprawia zdolność modelu do generalizacji
- Różne metody DA można łączyć ze sobą, co pozwala na tworzenie dowolnej liczby nowych instancji
- Trzy poziomy ogólności DA: przestrzeń cech, metody ogólne dla obrazów oraz metody specyficzne dla domeny
- Razem z ensemblingiem DA stanowi skuteczny sposób na podniesienie jakości predykcji
- Bardzo aktualnym zagadnieniem jest DA przy użyciu GANów
- Oprócz obrazów głównym obszarem zastosowania DA są sygnały (w szczególności dźwięki) oraz NLP (dużo mniej prac w tym temacie)

# Źródła

1. <https://sqream.com/where-will-the-big-be-spent-5-big-data-2015-predictions/>
2. <https://www.edureka.co/blog/interview-questions/deep-learning-interview-questions/>
3. <https://medium.com/nanonets/how-to-use-deep-learning-when-you-have-limited-data-part-2-data-augmentation-c26971dc8ced>
4. <https://medium.com/@prasad.pai/data-augmentation-techniques-in-cnn-using-tensorflow-371ae43d5be9>
5. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).

# Źródła

6. Wong, S. C., Gatt, A., Stamatescu, V., & McDonnell, M. D. (2016, November). Understanding data augmentation for classification: when to warp?. In *2016 international conference on digital image computing: techniques and applications (DICTA)* (pp. 1-6). IEEE.
7. <https://www.scilab.org/tutorials/computer-vision-%E2%80%93-image-transform>
8. <https://www.pling.org.uk/cs/cgvimg/shearing.png>
9. Zhong, Z., Zheng, L., Kang, G., Li, S., & Yang, Y. (2017). Random erasing data augmentation. *arXiv preprint arXiv:1708.04896*.
10. DeVries, T., & Taylor, G. W. (2017). Improved regularization of convolutional neural networks with cutout. *arXiv preprint arXiv:1708.04552*.

# Źródła

11. Li, Z., Guo, J., Jiao, W., Xu, P., Liu, B., & Zhao, X. (2018). Random linear interpolation data augmentation for person re-identification. *Multimedia Tools and Applications*, 1-17.
12. Buslaev, A., Parinov, A., Khvedchenya, E., Iglovikov, V. I., & Kalinin, A. A. (2018). Alumentations: fast and flexible image augmentations. *arXiv preprint arXiv:1809.06839*.
13. <https://iq.opengenus.org/autoencoder/>