



**POLITECHNIKA
RZESZOWSKA**
im. IGNACEGO ŁUKASIEWICZA



**WYDZIAŁ
ELEKTROTECHNIKI
I INFORMATYKI**
POLITECHNIKI RZESZOWSKIEJ



Augmentacja (rozszerzanie) treningowych zbiorów danych ukierunkowana na poprawę dokładności klasyfikacji

Wykład w ramach przedmiotu „Metody rozpoznawania obiektów i analizy ruchu”

Czym jest augmentacja zbiorów treningowych i po co się ją stosuje?

- ▶ Klasyfikatory wymagają dużej ilości danych treningowych do rozpoznawania obiektów z dużą dokładnością.
- ▶ Czasami nie mamy możliwości zapewnienia większej ilości próbek lub jest to uciążliwe. Przykładowo, aby utworzyć zbiór treningowy dla systemu rozpoznającego gesty musimy zaangażować wiele osób. Gesty powinny być zróżnicowane i reprezentatywne, a więc jedna/dwie osoby zazwyczaj nie wystarczą.
- ▶ Augmentacja polega na automatycznym generowaniu nowych próbek zbioru treningowego na podstawie już istniejących.
- ▶ Trenując klasyfikator na rozszerzonym w ten sposób zbiorze treningowym możemy zwiększyć dokładność rozpoznawania obiektów/gestów.

Sposoby augmentacji danych

- ▶ Rozszerzać możemy:
 - ▶ oryginalne dane wejściowe: np. obrazy, chmury punktów, szkielety lub ich sekwencje;
 - ▶ wyznaczone wcześniej wektory cech.

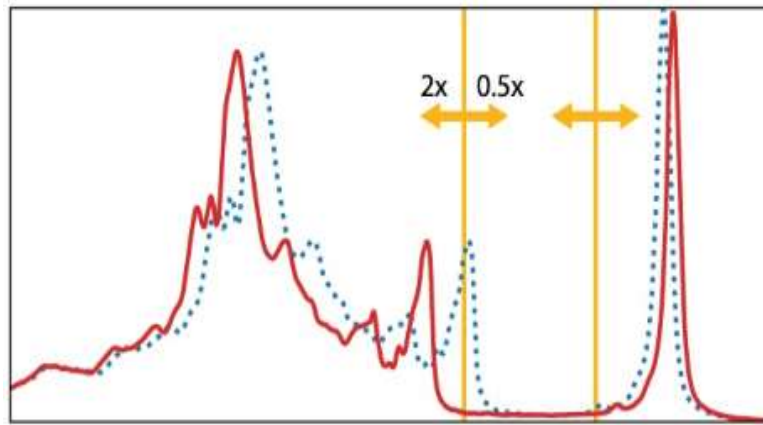
Augmentacja danych standardowych (niesekwencyjnych)

- ▶ Jeśli rozpoznajemy obiekty lub gesty statyczne, których próbki nie są sekwencjami, to augmentację wykonujemy najczęściej w następujący sposób:
 1. Augmentacja wektorów cech - wylosowanie kilku próbek należących do danej klasy, obliczenie średniej (arytmetycznej, geometrycznej, harmonicznnej itp.) każdej cechy i utworzenie nowej próbki na podstawie obliczonych średnich.
 2. Augmentacja chmur punktów – wylosowanie kilku chmur, następnie lekkie, losowe przemieszczenie w przestrzeni 3D wszystkich (lub części) punktów i utworzenie w ten sposób nowej chmury.
 3. Augmentacja obrazów:
 - ▶ Transformacje istniejących obrazów: obroty, odwracanie, zniekształcanie perspektywy, zmiany jasności i kontrastu.
 - ▶ Wygenerowanie nowych próbek przy użyciu generatywnych modeli uczenia, np. GAN (Generative Adversarial Network).

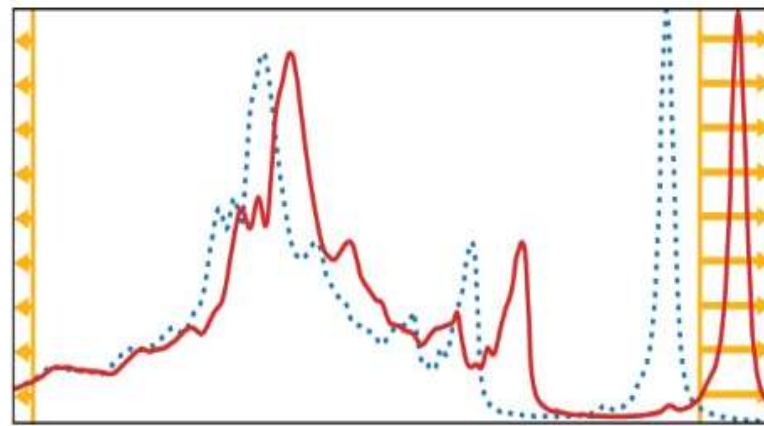
Augmentacja danych sekwencyjnych (szeregów czasowych)

- ▶ W przypadku rozpoznawania gestów dynamicznych, np. akcji wykonywanych przez ludzi albo innych danych o charakterze sekwencyjnym przykładowe metody augmentacji, to:
 1. Window Warping
 2. Window Slicing
 3. Enhanced Warped Windows
 4. Metody oparte o algorytm DTW - próbki są generowane korzystając ze ścieżek dopasowania wygenerowanych na podstawie dwóch przebiegów.

Window Warping i Window Slicing



Window Warping



Window Slicing

Źródło: <https://doi.org/10.1371/journal.pone.0254841.g001>

Efficient Warped Windows

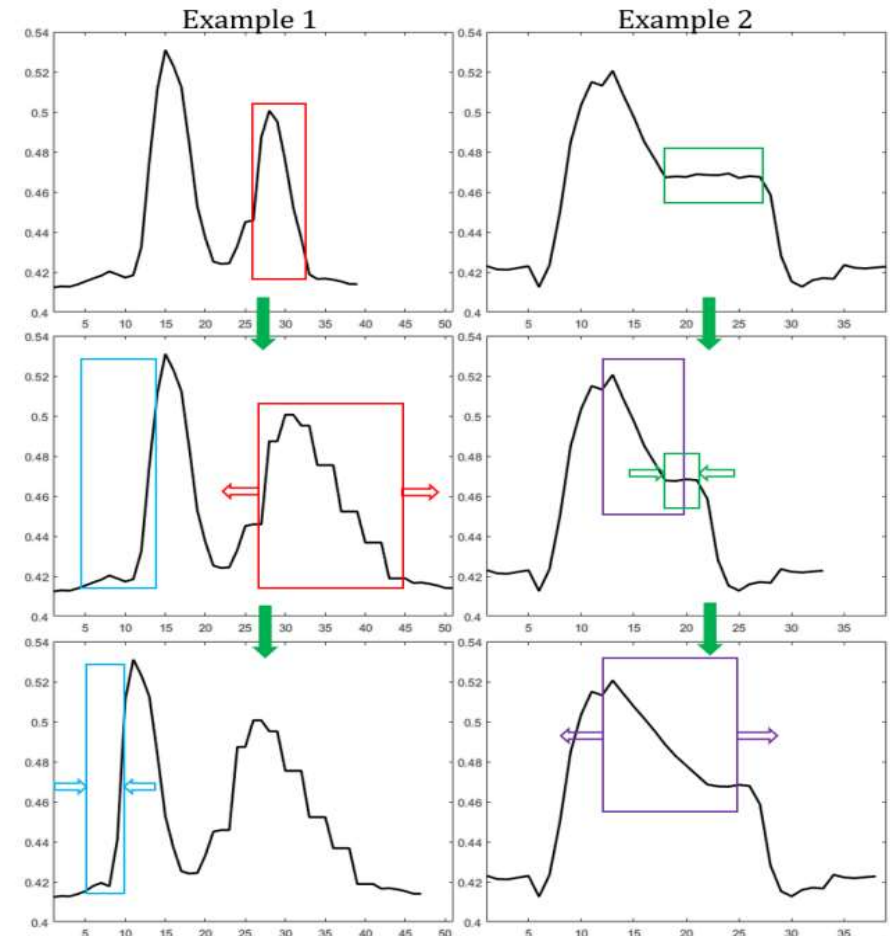
Efficient Warped Windows – rozszerzenie metody

Window Warping o elementy związane z interpolacją danych.

Efficient Warped Windows:

- ▶ losowo wybiera dwa wycinki;
- ▶ dla pierwszego z nich losuje jeden z dwóch operatorów:
 - ▶ nieliniowe rozciągnięcie czasu (znane z metody DTW) – **przykład 1**,
 - ▶ zwężenie (subsampling) – **przykład 2**;
- ▶ dla drugiego wycinka losuje jeden z następujących operatorów:
 - ▶ rozciągnięcie techniką interpolacji – **przykład 2**,
 - ▶ zwężenie (subsampling) – **przykład 1**.

Źródło: <https://doi.org/10.1016/j.procs.2022.09.360>



Eksperymenty – dokładność rozpoznawania akcji ludzkich po użyciu metod augmentacji

- ▶ Osiem popularnych zbiorów danych, klasyfikator BiLSTM
- ▶ W przypadku augmentacji zbiory zawierały dane oryginalne + dane sztucznie wygenerowane (nie same wygenerowane)

Metoda/ zbiór	MSRA I	MSRA II	MSRA III	UTD	UTK	FLORENCE	SYSU	KARD
BiLSTM								
Bez rozszerzania	68.8	66.9	81.3	72.4	91.4	79.9	54.9	78.7
WW	77.2	67.5	85.5	82.1	93.1	81.7	65.1	90.6
WS	78.2	68.5	84.3	80.7	93.1	80.2	65	90.6
EWW	78.8	71.2	86.3	85	93.5	82.1	64.5	91.9

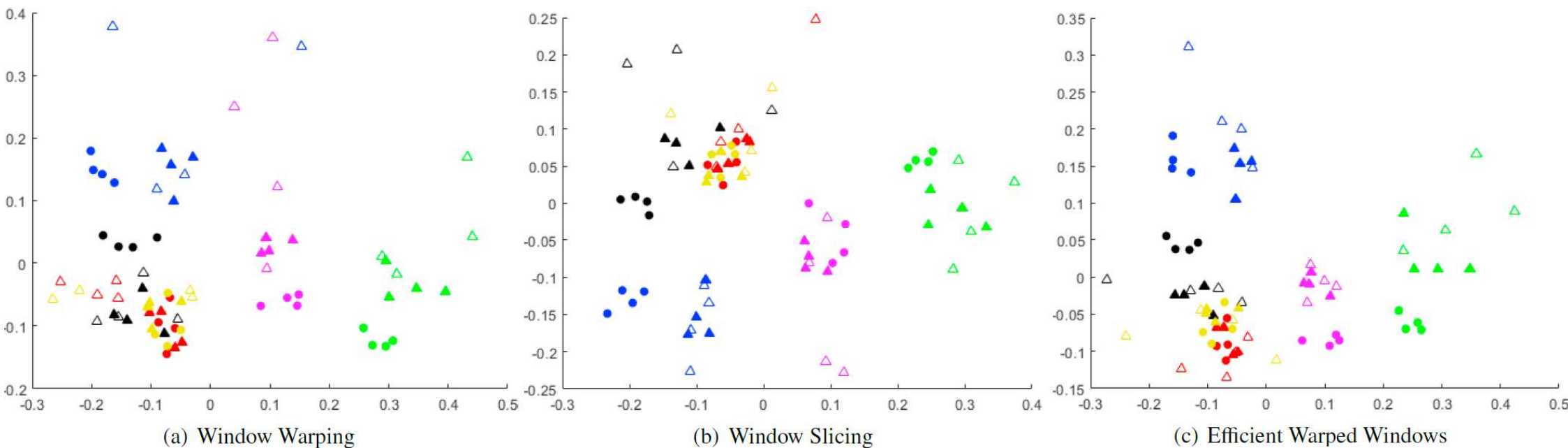
Eksperymenty – dokładność rozpoznawania akcji ludzkich po użyciu metod augmentacji

- ▶ Osiem popularnych zbiorów danych, klasyfikator k-najbliższych sąsiadów z odległością DTW
- ▶ W przypadku augmentacji zbiory zawierały dane oryginalne + dane sztucznie wygenerowane (nie same wygenerowane)

Metoda/ zbiór	MSRA I	MSRA II	MSRA III	UTD	UTK	FLORENCE	SYSU	KARD
kNN + DTW								
Bez rozszerzania	71.7	69	83.9	86.3	81.9	78.6	69.2	89.6
WW	70.6	69.7	84.2	86.3	80.7	78.4	67.2	90.9
WS	74.3	73.1	84	83.9	86.4	81.7	70.8	91.6
EWW	73.7	72	83.3	87	82.5	81	70.6	91.6

Eksperymenty – wizualne porównanie podobieństw sekwencji

- ▶ Reprezentacja sekwencji przy użyciu metody skalowania wielowymiarowego Kruskala (zbiór UTD-MHAD). Puste trójkąty oznaczają wygenerowane dane treningowe.



Przykładowy kod w Matlabie korzystający z metod augmentacji

```
%Wyznaczanie cech na podstawie danych sekwencyjnych  
%...  
  
[outTrain, outTrainLabels] = augmentation(train,trainLables);  
  
model = classifier.train(outTrain, outTrainLabels);  
resultClass = classifier.cassify(newSample);
```