

# Zastosowania DNN do analizy EEG

na podstawie:

♦ Human Brain Mapping 38:5391–5420 (2017) ♦

## Deep Learning With Convolutional Neural Networks for EEG Decoding and Visualization

Robin Tibor Schirrmeister <sup>1,2\*</sup> Jost Tobias Springenberg,<sup>2,3</sup>  
Lukas Dominique Josef Fiederer <sup>1,2,4</sup> Martin Glasstetter,<sup>1,2</sup>  
Katharina Eggenberger,<sup>2,5</sup> Michael Tangermann,<sup>2,6</sup> Frank Hutter,<sup>2,5</sup>  
Wolfram Burgard,<sup>2,7</sup> and Tonio Ball <sup>1,2</sup>

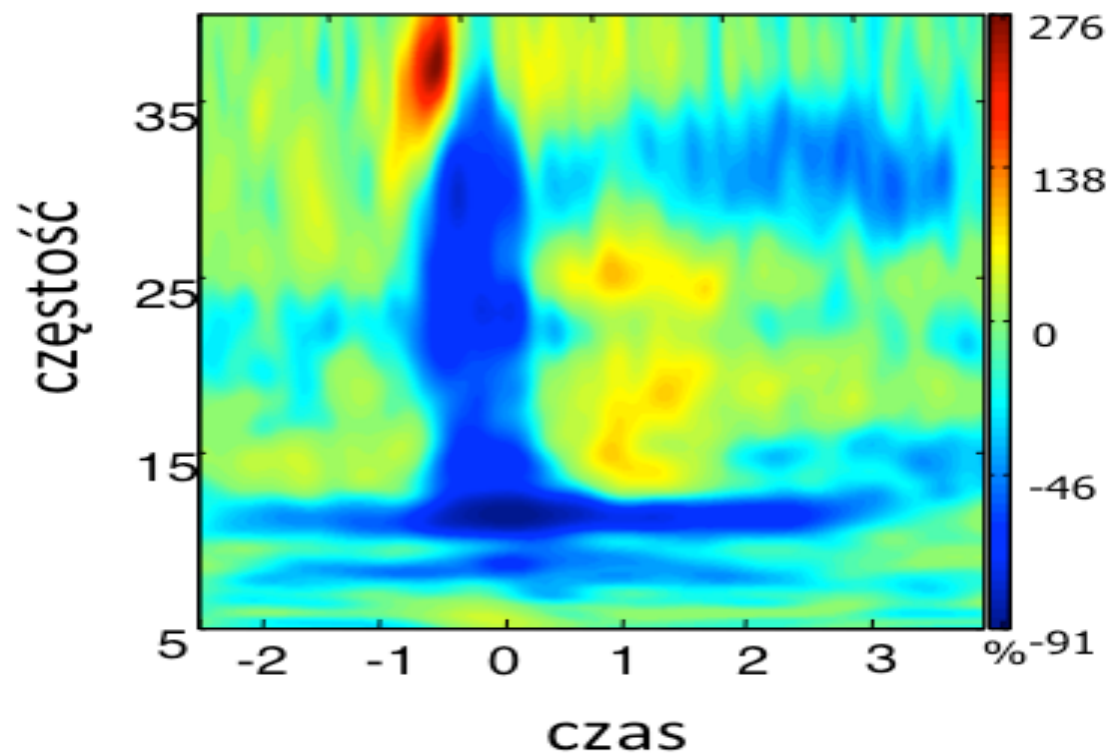
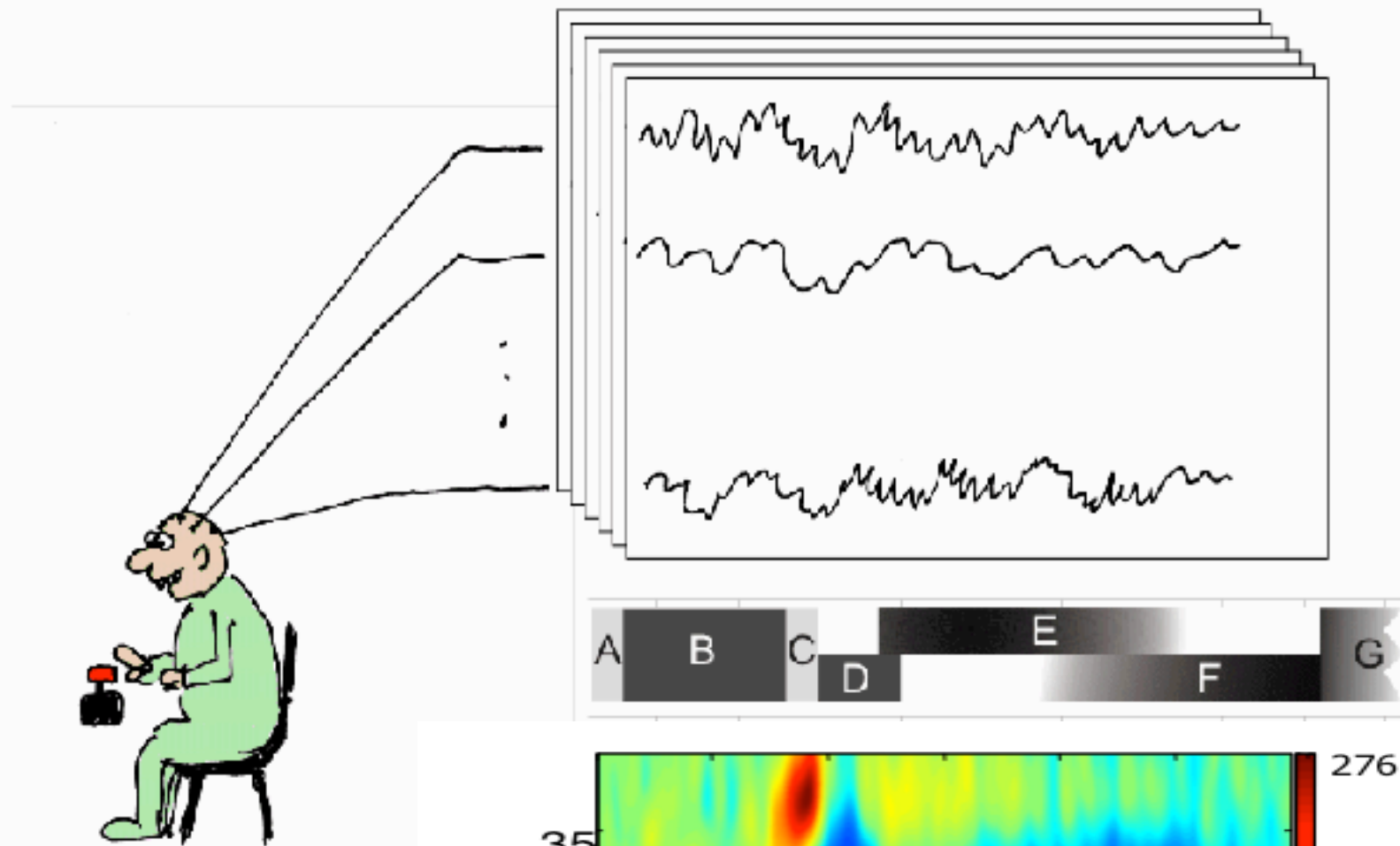
# Ogólnie

- Przykład zastosowania sieci głębokich do wytworzenia cech użytecznych do dekodowania sygnału EEG
- Demonstracja cech wytworzonych przez sieć
- Porównanie do cech opracowanych w tej dziedzinie przez ekspertów

# Reprezentacja danych wejściowych

- Sygnał EEG jest zazwyczaj wielokanałowy
- Kanały nie są w pełni niezależne - mają strukturę korelacyjną w przestrzeni
- W ramach danego kanału występuje struktura korelacji czasowych
- Istotne informacje niesione są też przez różne pasma częstotliwości
- W klasycznym podejściu wykorzystujemy filtry przestrzenne, filtry częstotliwościowe, analizę morfologiczną (kształtu)

# Paradygmat eksperymentów ERD/ERS

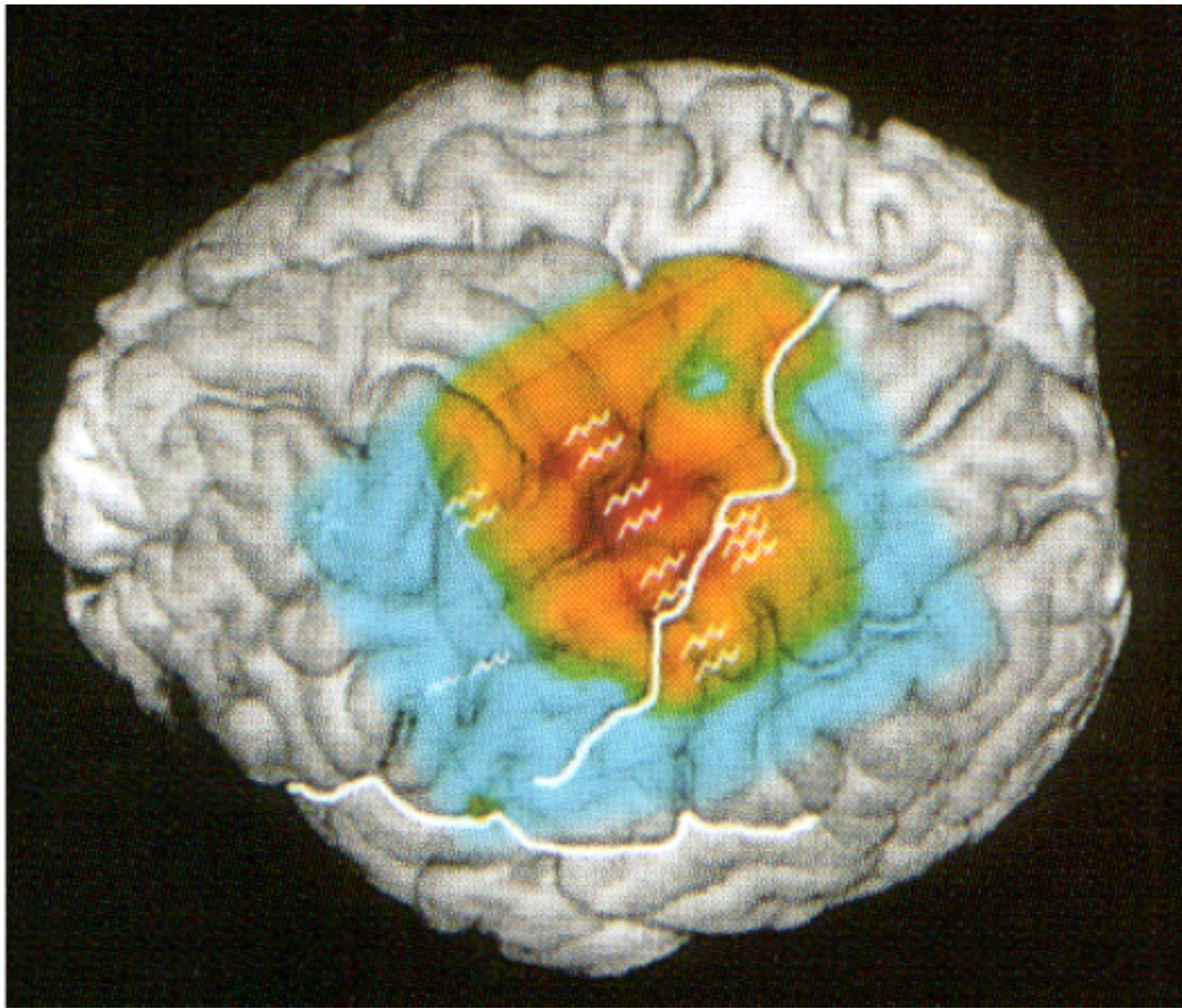


Przykładowa mapa zjawiska synchronizacji i desynchronizacji sygnału EEG w przestrzeni czas-częstość dla sygnału zarejestrowanego przez elektrodę C3.

Osoba badana wykonywała szybki ruch palcem w momencie oznaczonym 0.

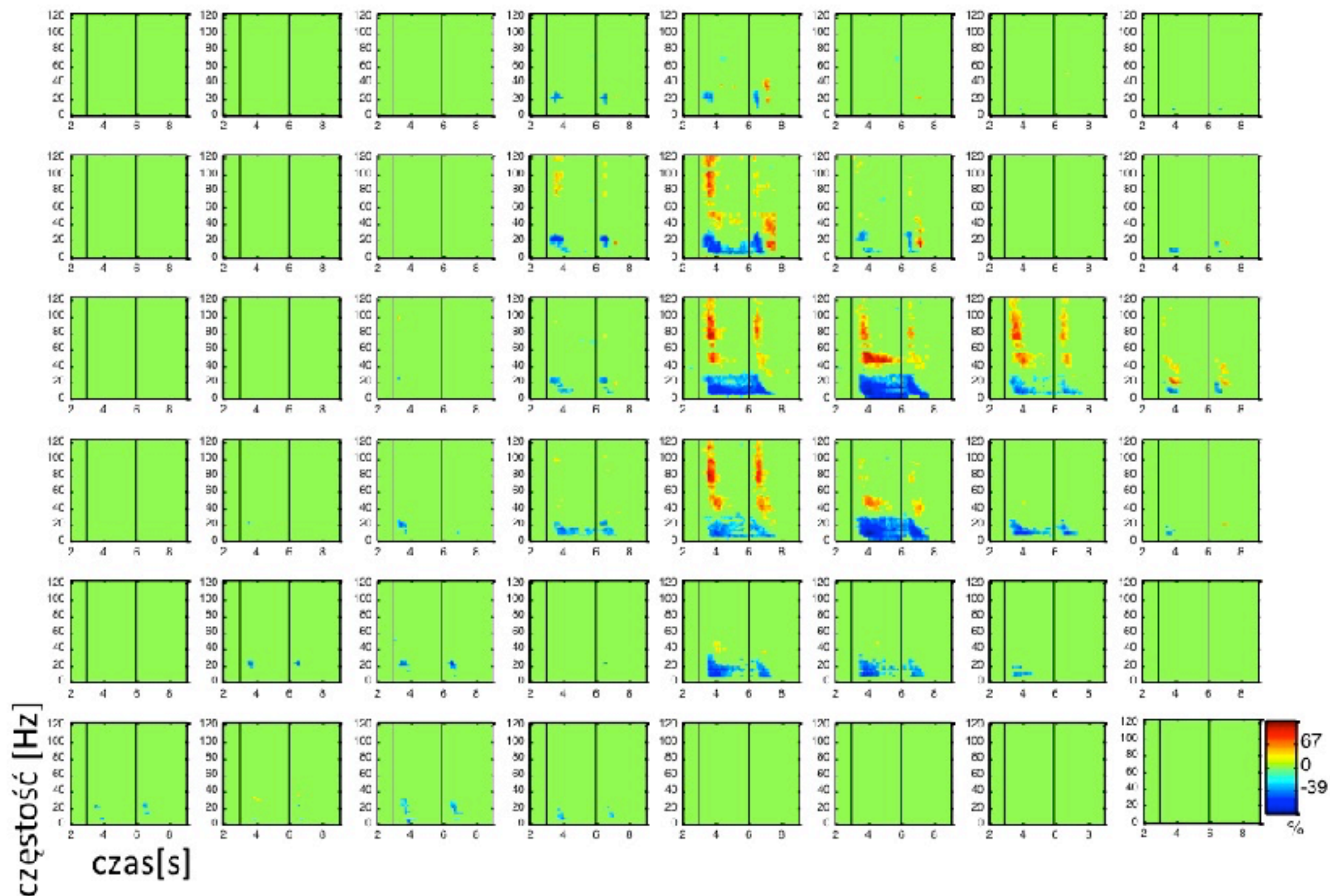
Kolory odpowiadają procentowej zmianie mocy względem poprzedzającego okresu 2s, zgodnie ze skalą barw umieszczoną po prawej stronie. Na poziomej osi - czas w s, na pionowej częstość w Hz.





**Mapowanie kory czuciowo-ruchowej za pomocą ERD/ERS. Pacjent z przyczyn medycznych miał umieszczoną na korze motorycznej siatkę elektrod (6x8, odstęp 1cm). Pacjent miał zaciskać pięść w odpowiedzi na bodziec wzrokowy przez cały czas trwania bodźca (3s) Prezentowane mapy istotnych statystycznie zmian uzyskano na podstawie 49 wolnych od artefaktów realizacji.**

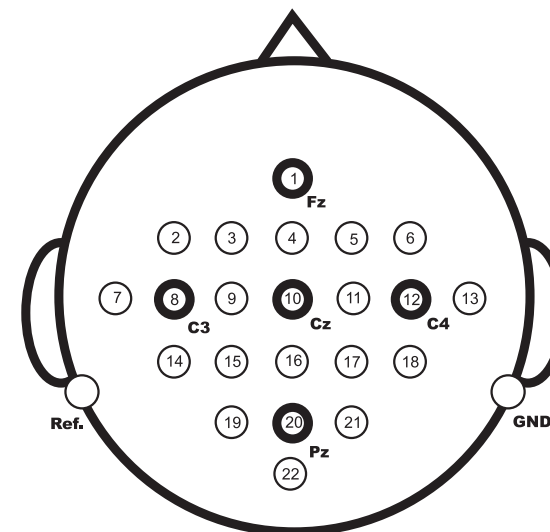
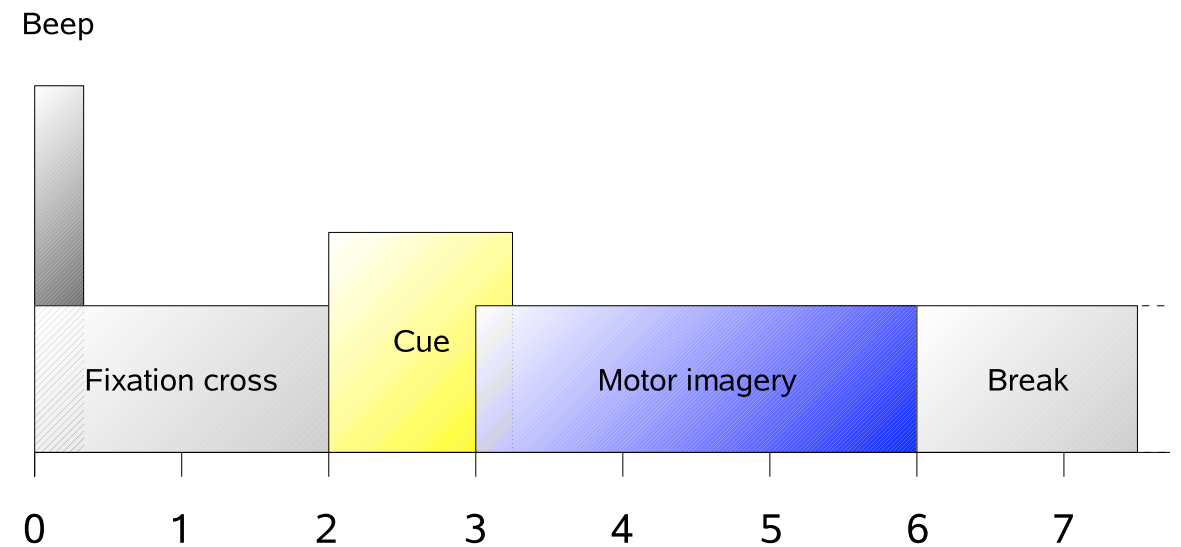




**Mapowanie kory czuciowo-ruchowej za pomocą ERD/ERS. Pacjent z przyczyn medycznych miał umieszczoną na korze motorycznej siatkę elektrod (6x8, odstęp 1cm). Pacjent miał zaciskać pięść w odpowiedzi na bodziec wzrokowy przez cały czas trwania bodźca (3s) Prezentowane mapy istotnych statystycznie zmian uzyskano na podstawie 49 wolnych od artefaktów realizacji.**

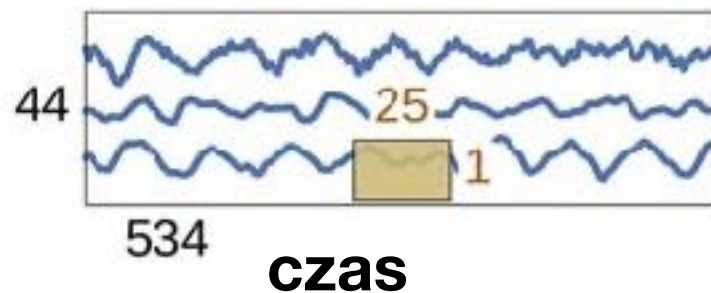
# Przykładowe dane: BCI competition IV dataset 2a

- wskazane wyobrażenia ruchowe ręka lewa albo prawa, stopy, język
- 22 kanały EEG
  - filtrowanie: 0.5-100Hz; filtr sieciowy),
- 250Hz próbkowanie,
- 4 klasy,
- 9 osób
- 288 epok danych na osobę



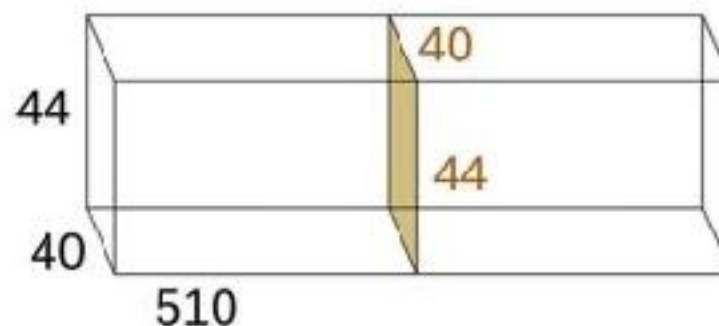
# Architektury

**filtry  
czasowe**  
40 Units



dostajemy 44 sygnały,  
każdy przefiltrowany czasowo  
na 40 sposobów

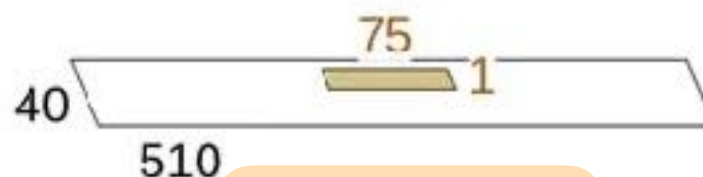
**filtry  
przestrzenne**  
40 Units



dostajemy 40 różnych  
liniowych kombinacji  
44 (syg.) x 40 (filt. czasowych)

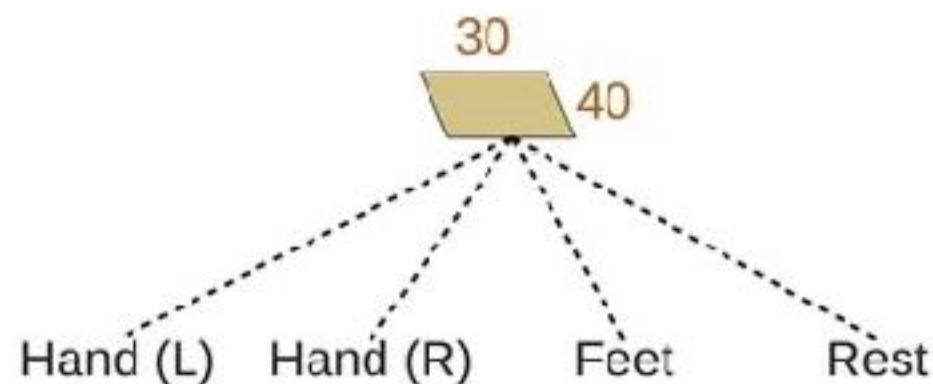
Square

**Mean Pooling**  
Stride 15x1



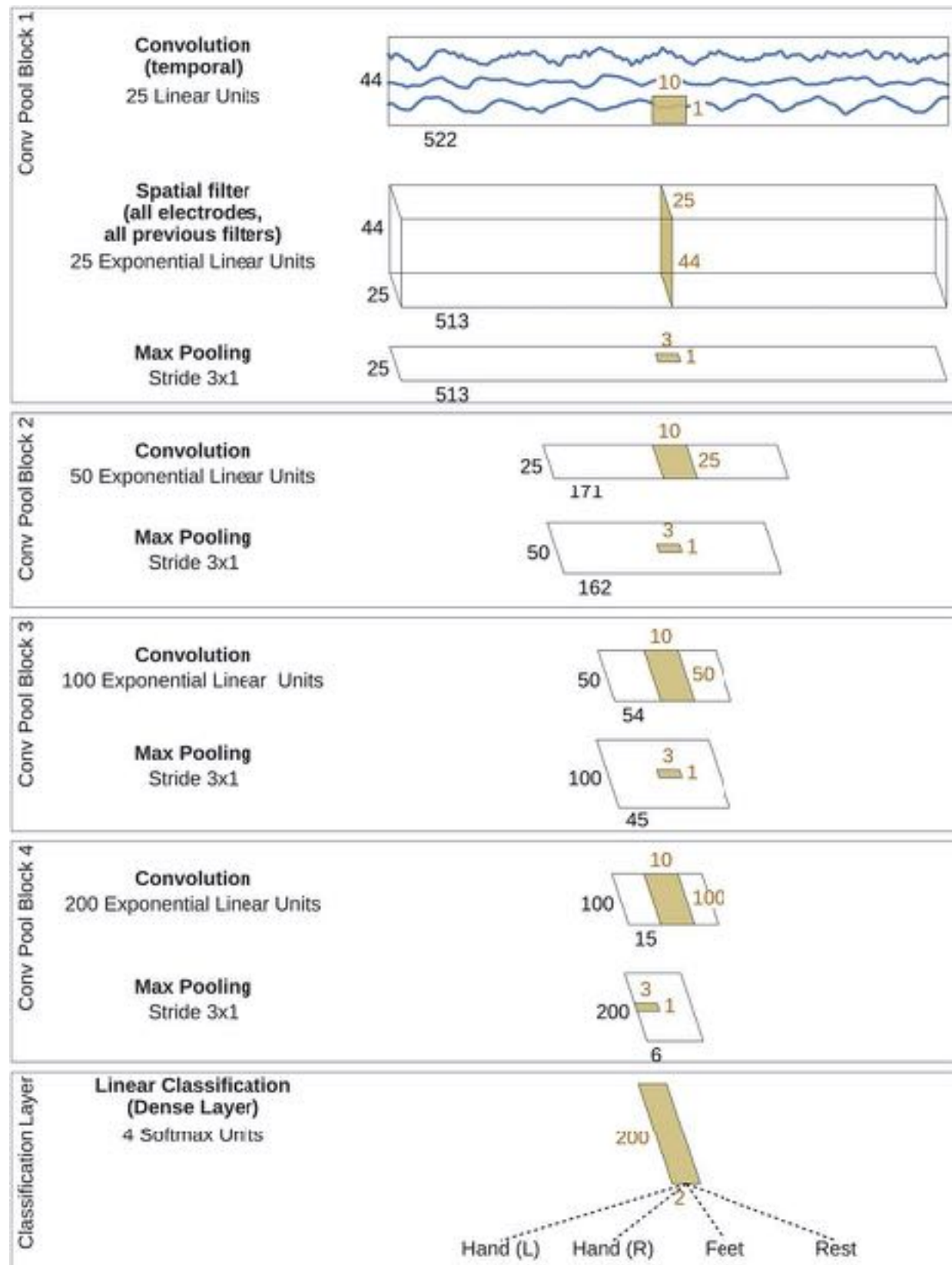
Log

**Linear Classification  
(Dense Layer+Softmax)**  
4 Units





# Architektury



**Jednostki typu ELU**

$$f(x) = \begin{cases} x & \text{dla } x \geq 0 \\ e^x - 1 & \text{dla } x < 0 \end{cases}$$

# Koncepcja pola recepcyjnego

Classification Layer  
Predictions

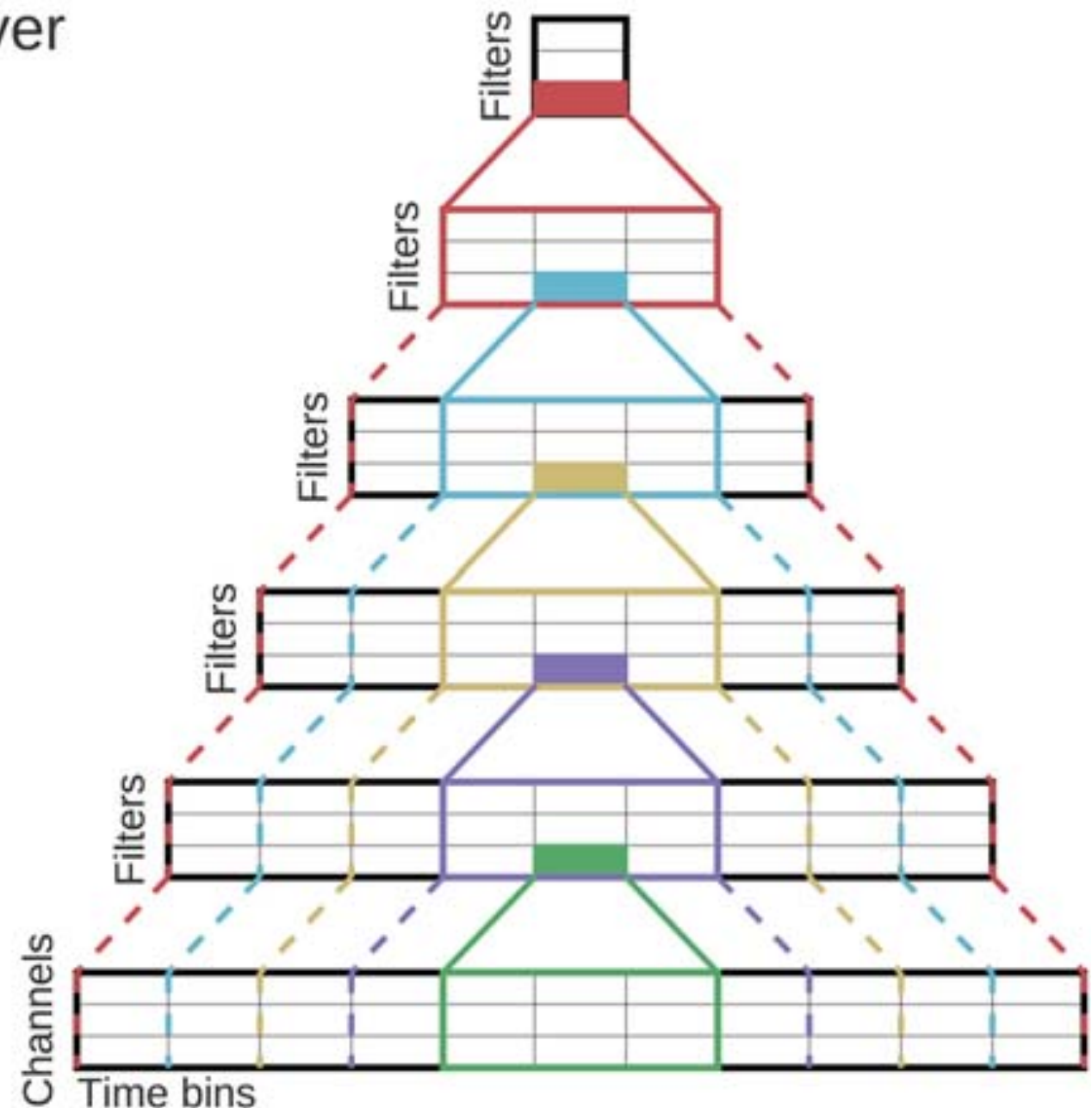
Layer 4  
Unit Output

Layer 3  
Unit Output

Layer 2  
Unit Output

Layer 1  
Unit Output

Input  
EEG Channels



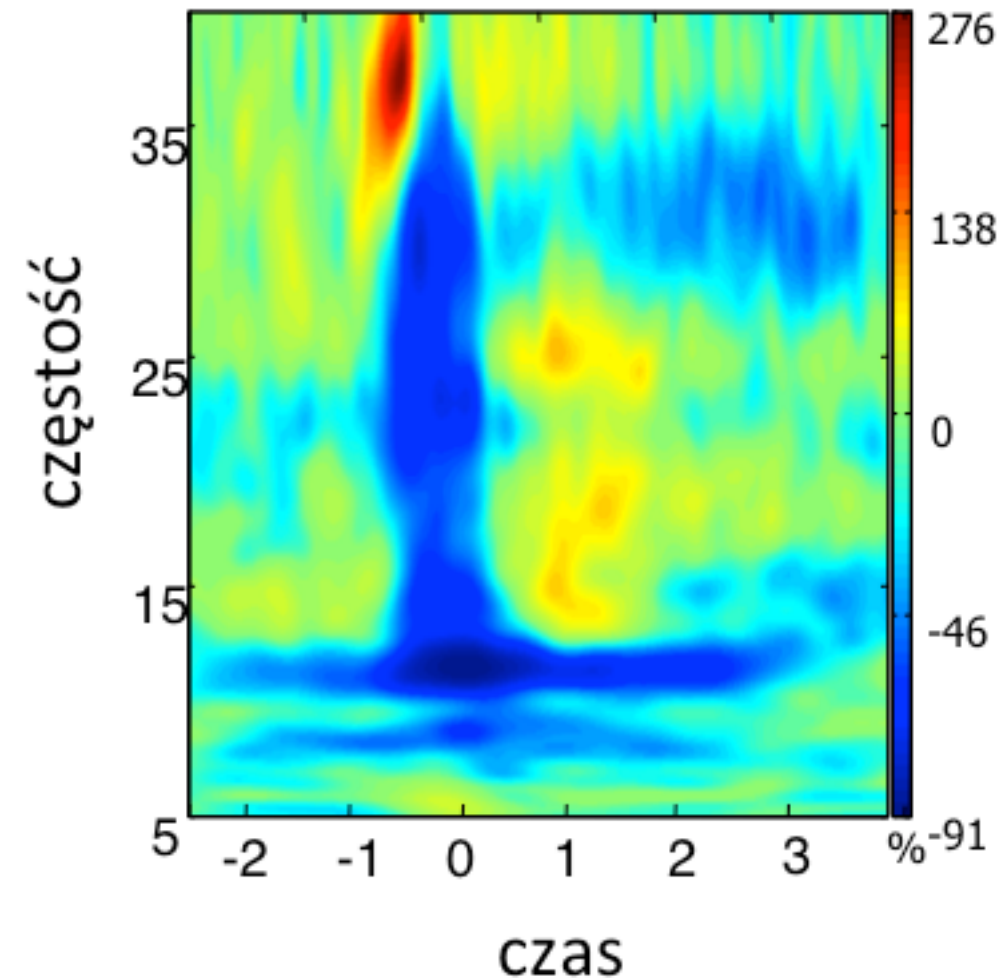
- wypełniane prostokąty to pojedyncze jednostki
- linie ciągłe to ich bezpośrednie pole recepcyjne
- linie przerywane to to efektywne pole recepcyjne - sygnał z tego pola w różnym stopniu przetworzony, dociera do danej jednostki

# Wgląd w wyuczone cechy

- mapy korelacji pomiędzy wyjściem jednostek, a wartościami cech wejściowych.
  - Jako cechy wejściowe brane były średnia wartość mocy w pewnym paśmie częstości przypadająca na pole recepcyjne danej jednostki.
- mapy korelacji perturbacji wartości wejściowych i predykcjami sieci

# Wgląd w wyuczone cechy

- mapy korelacji pomiędzy wyjściem jednostek, a wartościami cech wejściowych.
- Jako cechy wejściowe brane były średnia wartość mocy w pewnym paśmie częstości przypadająca na pole recepcyjne danej jednostki.
- mapy korelacji perturbacji wartości wejściowych i predykcjami sieci

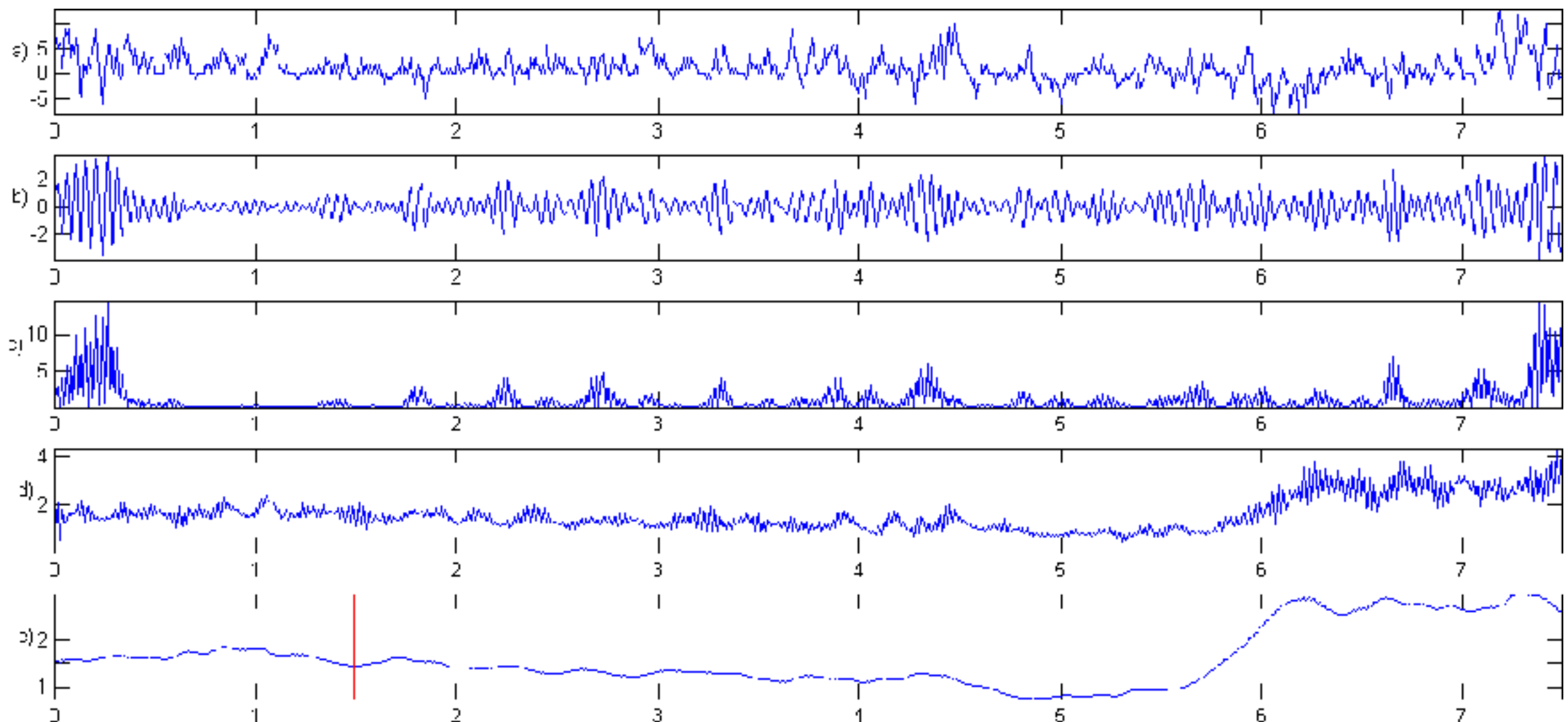


Przykładowa mapa zjawiska synchronizacji i desynchronizacji sygnału EEG w przestrzeni czas-częstość dla sygnału zarejestrowanego przez elektrodę C3. Osoba badana wykonywała szybki ruch palcem w momencie oznaczonym 0. Kolory odpowiadają procentowej zmianie mocy względem poprzedzającego okresu 2s, zgodnie ze skalą barw umieszczoną po prawej stronie. Na poziomej osi - czas w s, na pionowej częstość w Hz.



# Mapy korelacji pomiędzy wyjściami jednostek, a wartościami cech wejściowych.

- dla każdego interesującego pasma częstotliwości sygnał był filtrowany pasmowo i była obliczana obwiednia
- dla każdego pasma i dla każdego pola recepcyjnego danej warstwy obliczono średni kwadrat obwiedni



# Mapy korelacji pomiędzy wyjściami jednostek, a wartościami cech wejściowych.

- obliczono średnią moc w poszczególnych interesujących pasmach częstości dla pól recepcyjnych. Tu przykładowe trzy pasma w 10 próbach dla wybranych pól recepcyjnych.



# Mapy korelacji pomiędzy wyjściami jednostek, a wartościami cech wejściowych.

- obliczono średnią moc w poszczególnych interesujących pasmach częstości dla pól recepcyjnych. Tu przykładowe trzy pasma w 10 próbach dla wybranych pól recepcyjnych.



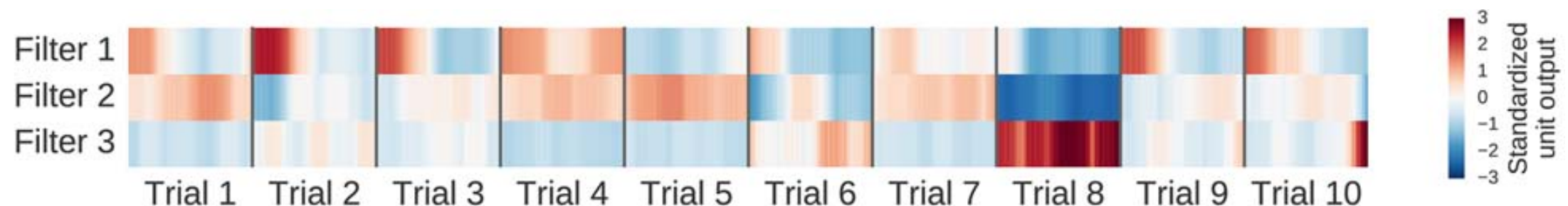
- obliczono odpowiedzi jednostek dla oryginalnego sygnału dla tych prób.

# Mapy korelacji pomiędzy wyjściami jednostek, a wartościami cech wejściowych.

- obliczono średnią moc w poszczególnych interesujących pasmach częstości dla pól recepcyjnych. Tu przykładowe trzy pasma w 10 próbach dla wybranych pól recepcyjnych.



- obliczono odpowiedzi jednostek dla oryginalnego sygnału dla tych prób.



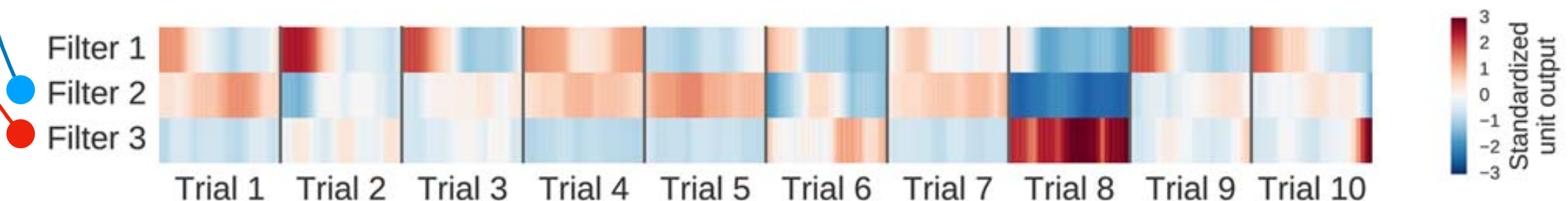


# Mapy korelacji pomiędzy wyjściami jednostek, a wartościami cech wejściowych.

- obliczono średnią moc w poszczególnych interesujących pasmach częstości dla pól recepcyjnych. Tu przykładowe trzy pasma w 10 próbach dla wybranych pól recepcyjnych.

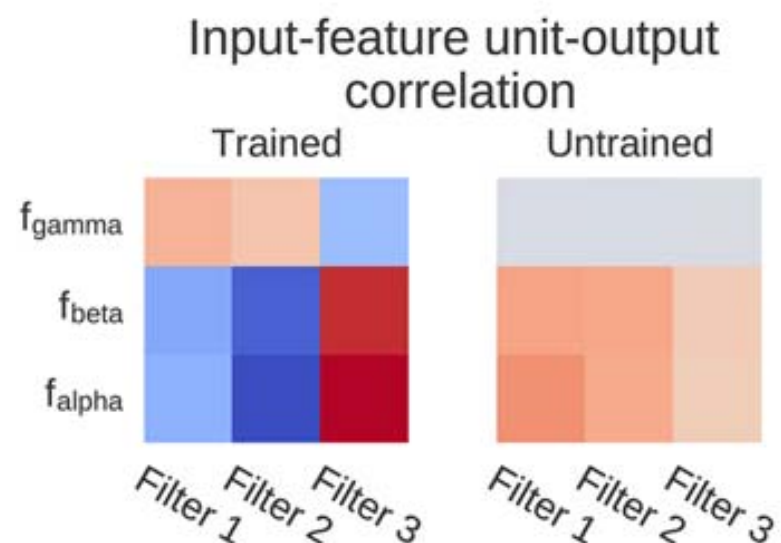


- obliczono odpowiedzi jednostek dla oryginalnego sygnału dla tych prób.

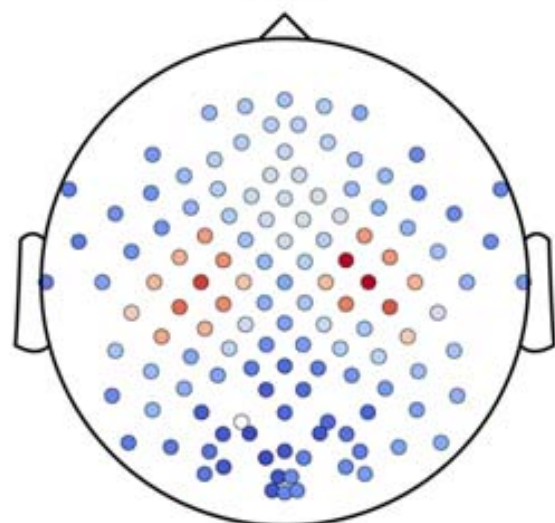
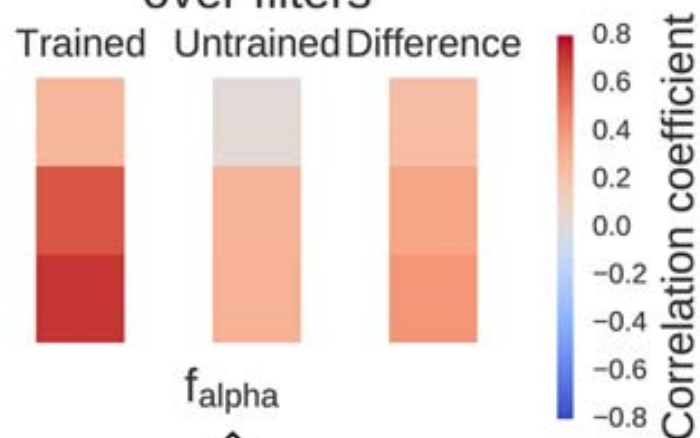


- w tym przykładzie Filter 2 ma odwrotne zachowanie, a Filter 3 zachowuje się podobnie jak średnia moc w paśmie alfa w jego polu recepcyjnym

# Mapy korelacji pomiędzy wyjściami jednostek, a wartościami cech wejściowych.



Mean of absolute correlation over filters



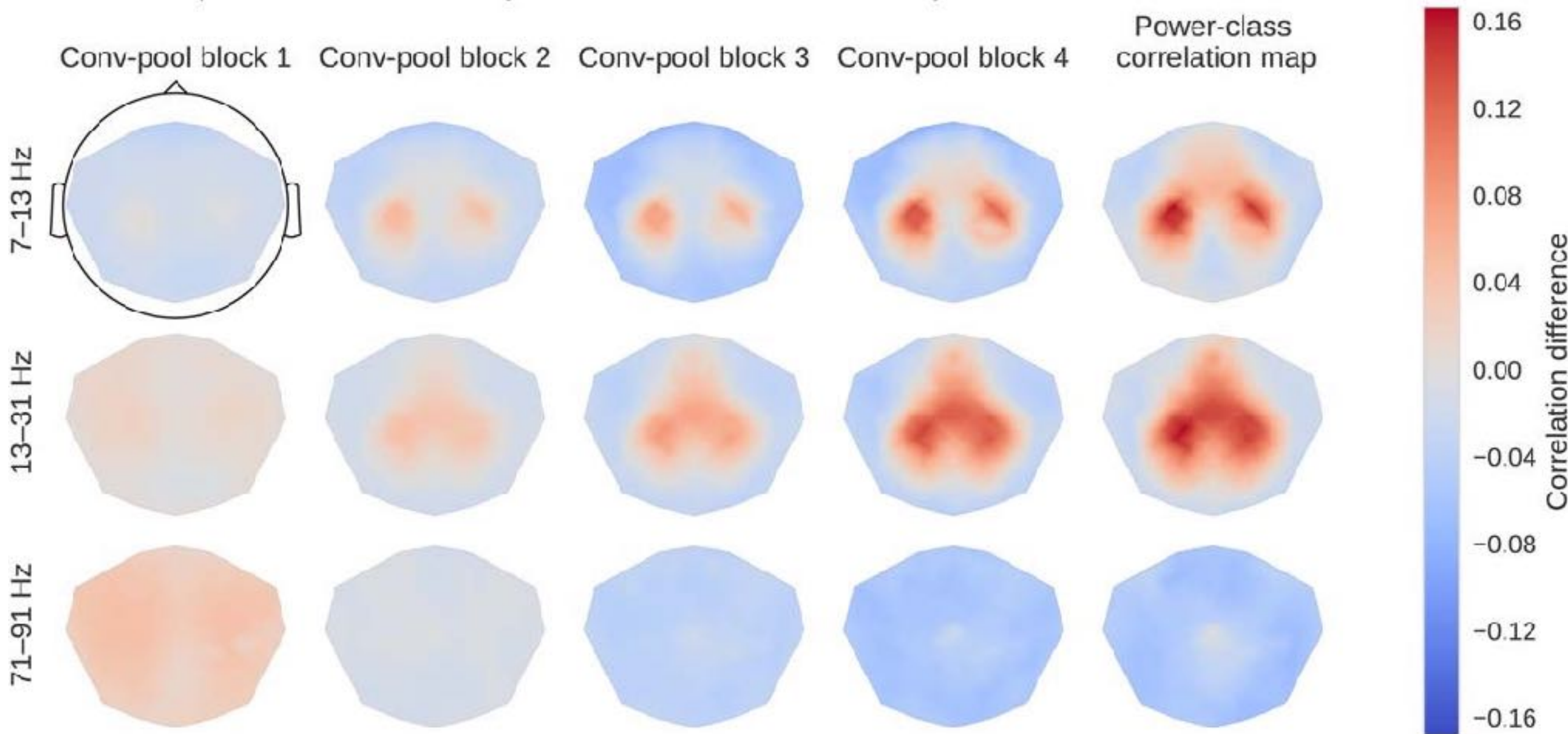
- obliczono korelacje pomiędzy średnimi kwadratami obwiedni w poszczególnych polach recepcyjnych a odpowiedzią jednostek (filtrów) -> to powinno ilustrować czy dana jednostka mogłaby realizować filtrowanie w danej częstotliwości
- zrobiono to dla sieci po i przed treningiem
- uśredniono wartości bezwzględne odpowiedzi jednostek
- na rysunku topograficznym zilustrowano rozkład przestrzenny różnic w korelacjach dla danego pasma przed i po treningu

# Poszukiwanie nieznanych cech dyskryminujących klasy

- skorelowano dane wyjściowe jednostek dla każdej warstwy bezpośrednio z etykietami klas.
- Następnie obliczono pojedynczy bezwzględny współczynnik korelacji na warstwę na dwa sposoby:
  - obliczono średni bezwzględny współczynnik korelacji dla wszystkich klas i wszystkich filtrów. Korelacje te powinny pokazywać, jak silnie dane wyjściowe filtra kodują etykiety klas.
  - obliczono maksymalne bezwzględne współczynniki korelacji dla każdej klasy dla wszystkich filtrów, a następnie uśredniono maksima z czterech klas. Korelacje te powinny pokazywać, jak mocno dane wyjściowe jednostki kodują etykiety klas dla najbardziej „informatywnych” filtrów.
- Wreszcie, w przypadku obu wersji porównano różnicę tych korelacji między modelem wytrenowanym a niewytrenowanym.
- To podejście pozwoliło pokazać, w jaki sposób powstają korelacje klasy w kolejnych warstwach ConvNet.

# Poszukiwanie nieznanych cech dyskryminujących klasy

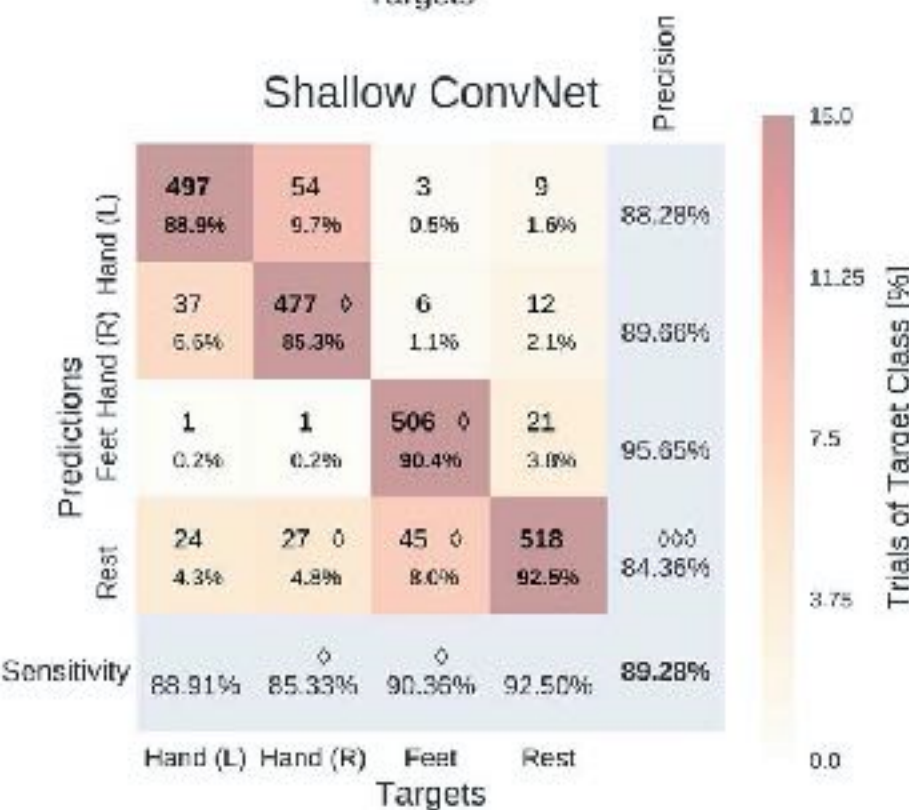
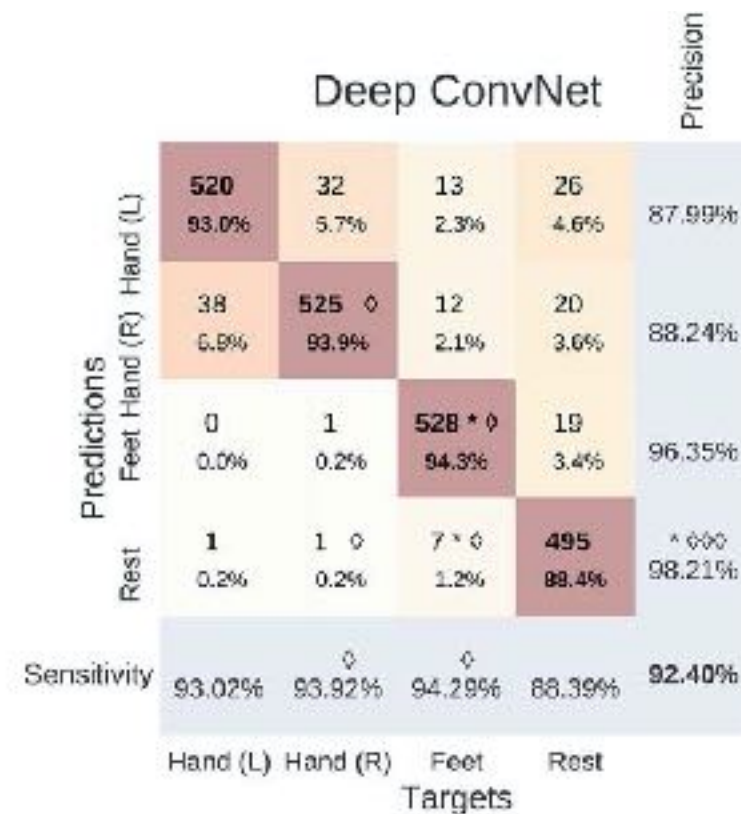
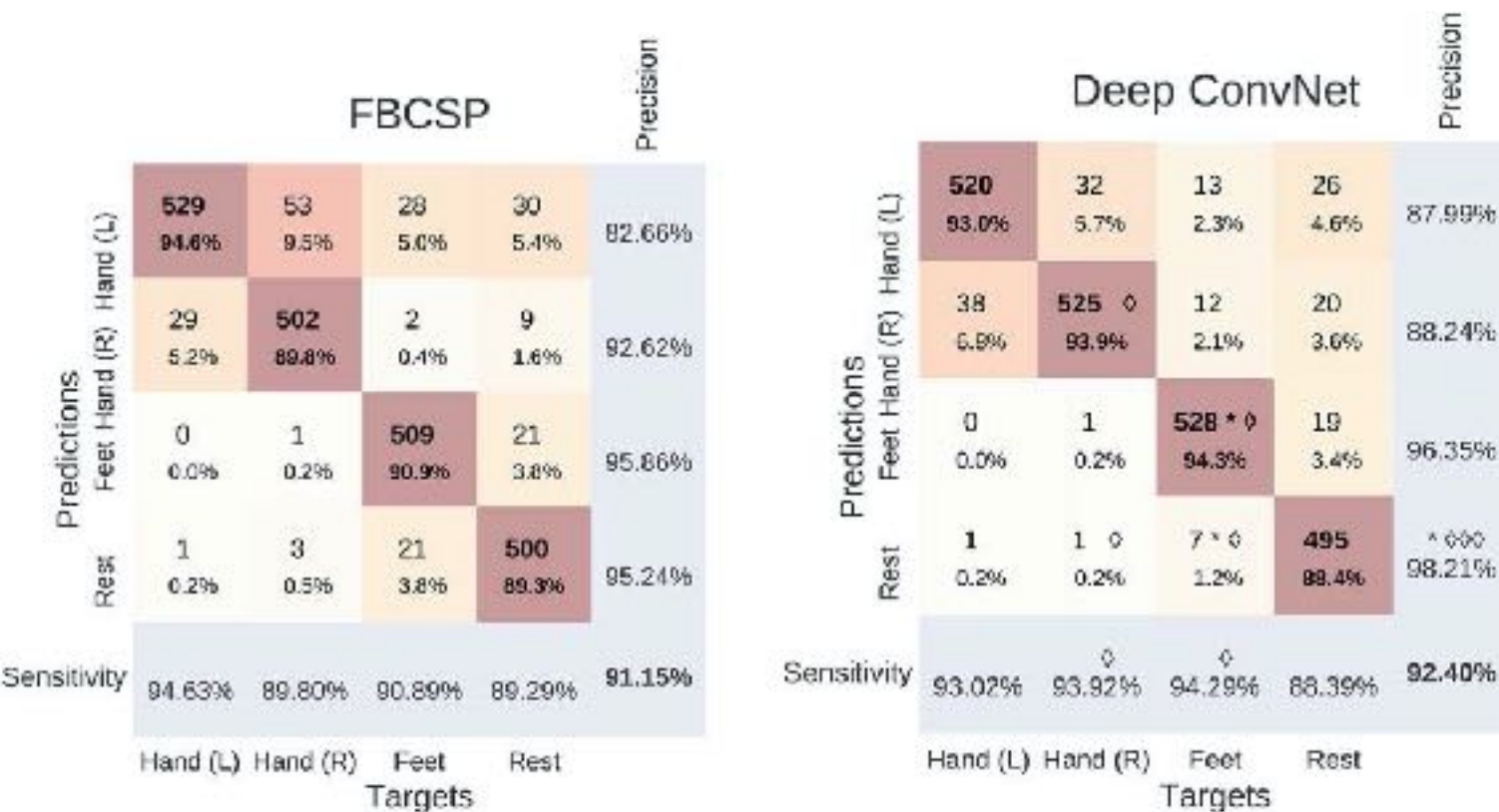
Input-feature unit-output network correlation maps



- To podejście pozwoliło pokazać, w jaki sposób powstają korelacje klasy w kolejnych warstwach ConvNet.



# Podsumowanie wyników



- Zastosowanie dropoutu, batch-normalizacji i jednostek ELU pozwoliło osiągnąć wyniki porównywalne z najlepszymi rozwiązaniami opartymi na klasycznych filtrach przestrzennych i pasmowych.
- Jak sugerują w konkluzjach Autorzy zastosowanie LRP być może pomogłoby w odkryciu cech wydobytych przez sieć, innych niż tylko związki z topografią i pasmami częstości.