

**AGH**

**AKADEMIA GÓRNICZO-HUTNICZA  
IM. STANISŁAWA STASZICA W KRAKOWIE**

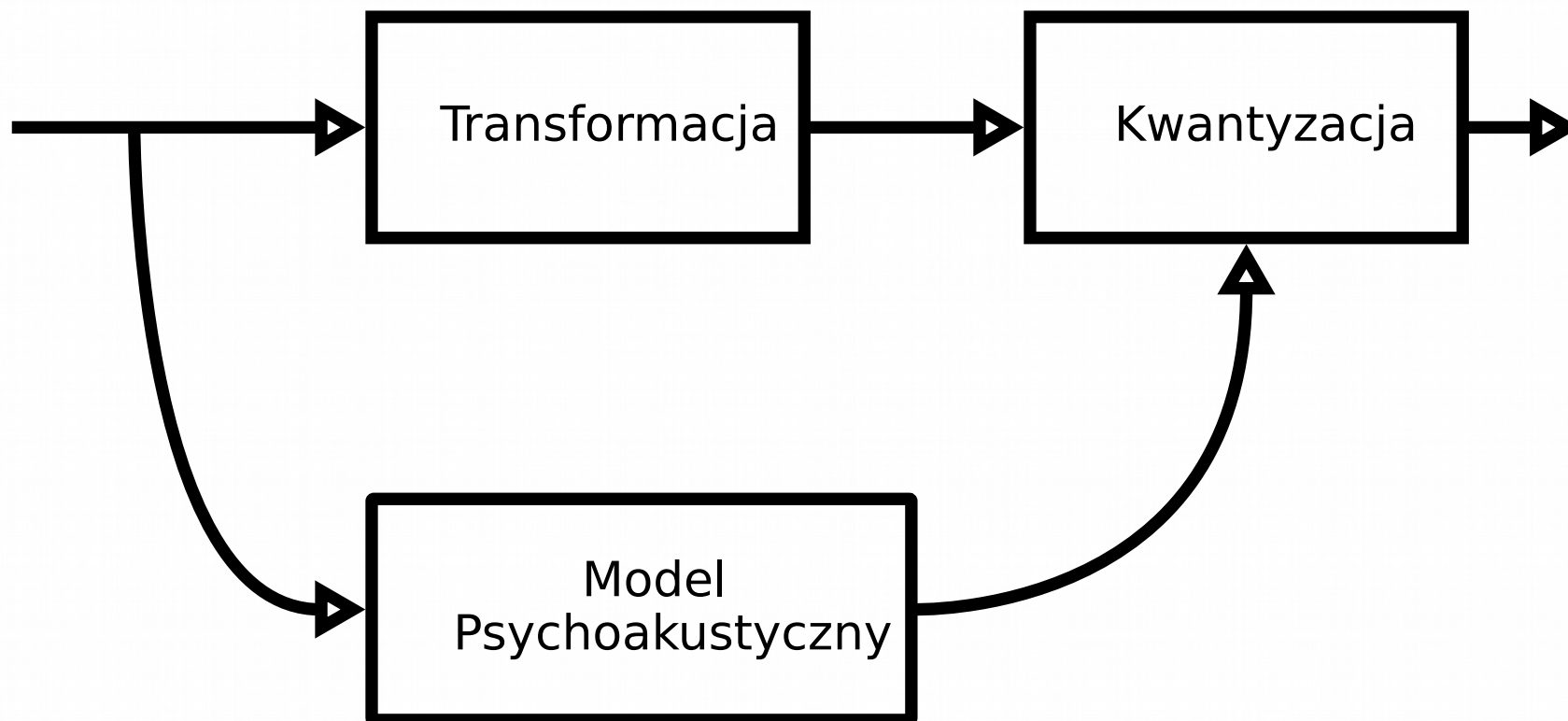
**AGH UNIVERSITY OF SCIENCE  
AND TECHNOLOGY**

# Zastosowanie uczenia maszynowego do wykrycia metody kompresji stratnej sygnału audio

Szymon Mikulicz

Data: 27.06.2019 r.

# Kompresja stratna audio



# Model Psychoakustyczny

- » Algorytm dobierający parametry analizy sygnału oraz jego kwantyzacji w sposób jak najmniej zauważalny dla ludzkiego ucha
- » Wyznaczany eksperymentalnie na podstawie testów odsłuchowych i wiedzy z zakresu psychoakustyki

# STFT

- » Algorytm pozwalający na analizę sygnału pod względem zarówno czasu jak i częstotliwości poprzez stosowanie transformaty Fouriera na kolejnych fragmentach sygnału
- » Fragmenty sygnału mnożone są przez tzw. okno w celu zniwelowania efektów brzegowych

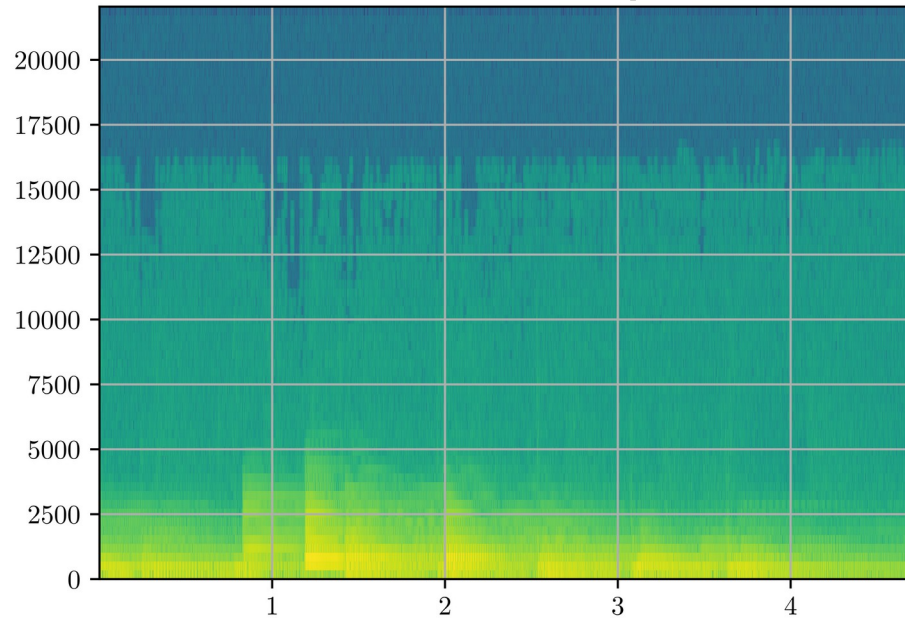


# Kompresja stratna audio

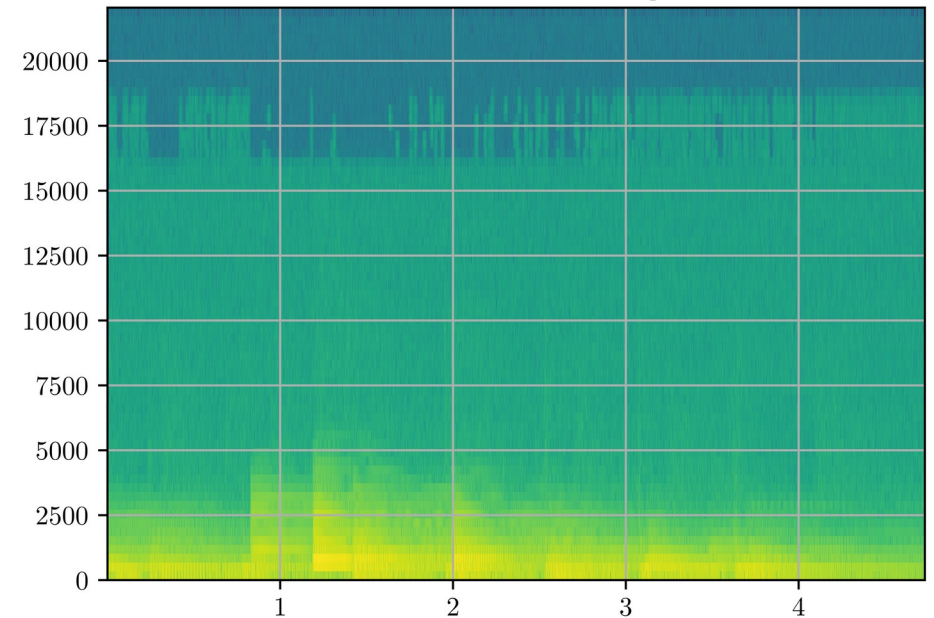


AGH

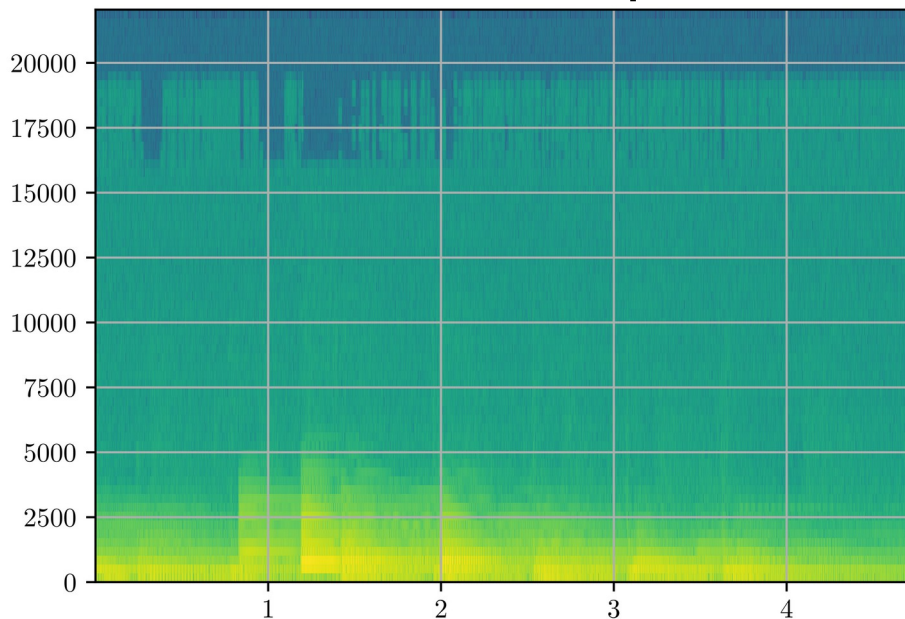
MP3 128 kbps



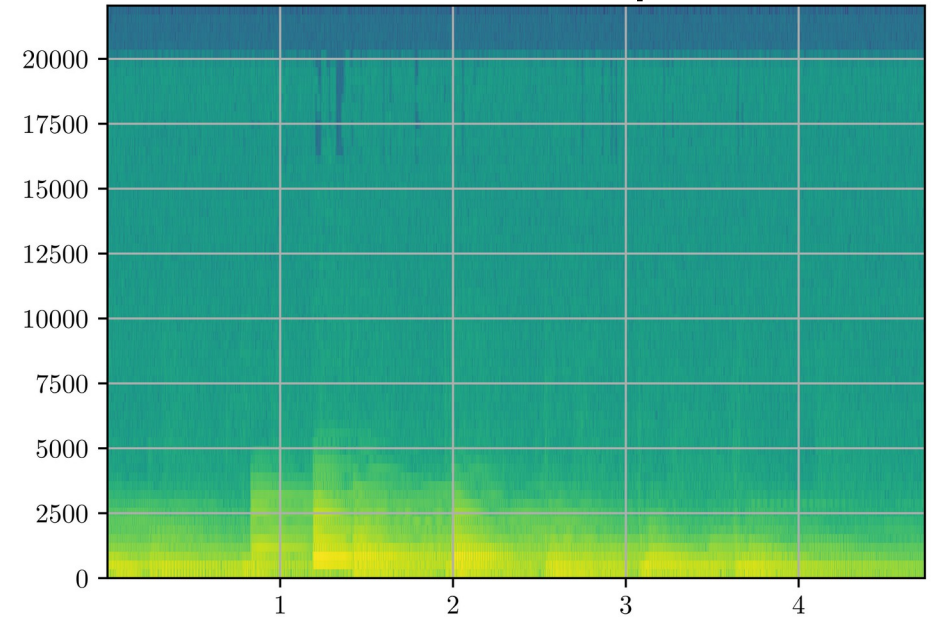
MP3 196 kbps



MP3 256 kbps



MP3 320 kbps



# Dane

- » 8192 STFT z 30-tosekundowych fragmentów sygnałów audio (muzyki)
- » Parametry STFT: długość okna 512, skok 256, okno hanninga
- » 75% przeznaczone na dane treningowe, 25% przeznaczone na dane testowe
- » Każdy z fragmentów został poddany kompresji MP3 128kbps, 192kbps, 320kbps lub nie został

# Wykorzystane oprogramowanie

- » Język programowania Julia
- » Biblioteka do uczenia maszynowego:  
Knet (Koç University deep learning  
framework)





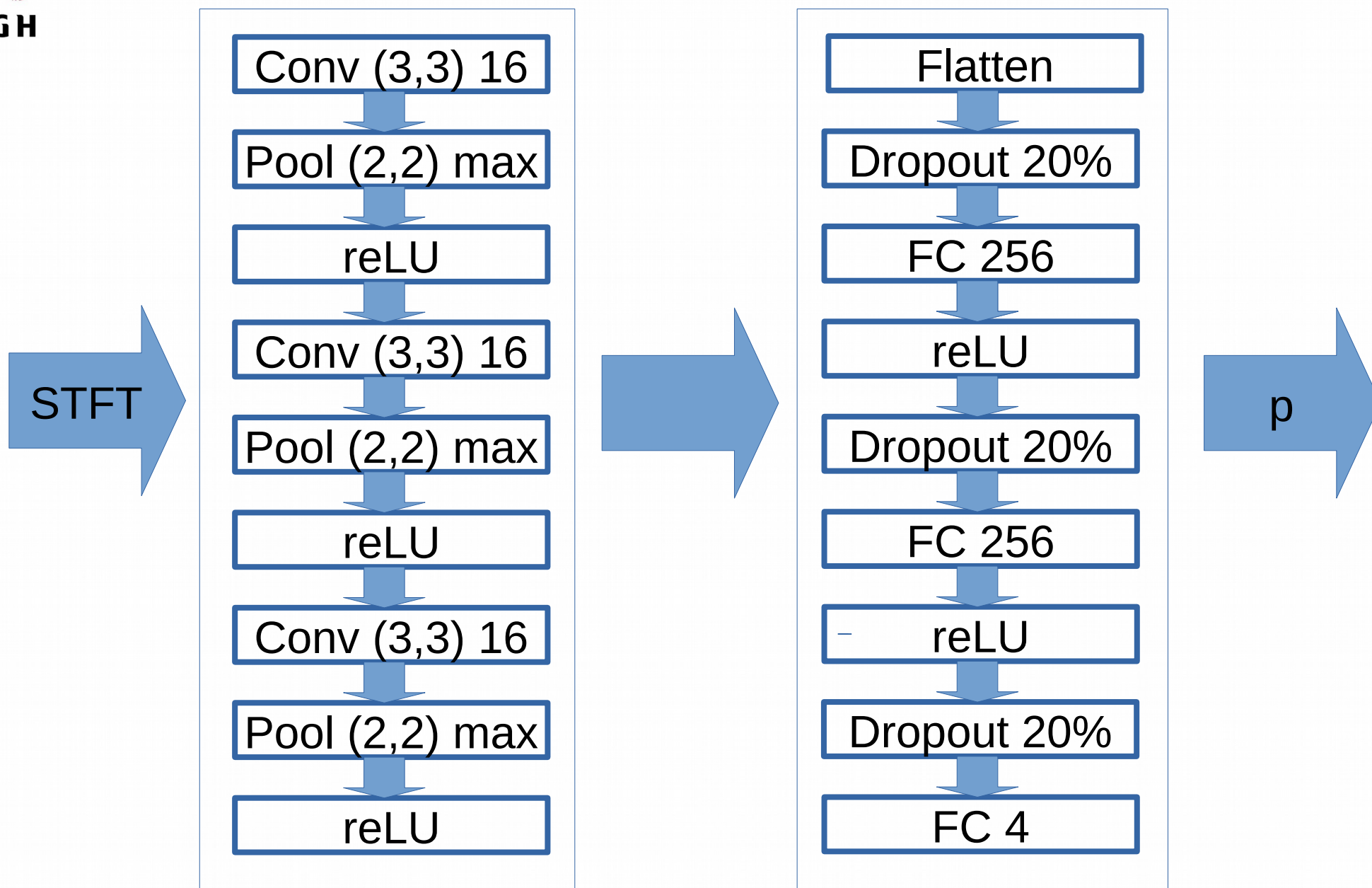
- » Język dynamiczny wysokiego poziomu
- » Zaprojektowany do wysokowydajnościowych obliczeń numerycznych
- » Wykorzystuje JIT (Just-In-Time Compilation) do uzyskania przewagi wydajnościowej nad głównymi rywalami (Python, R, GNU Octave)



# Biblioteka Knet

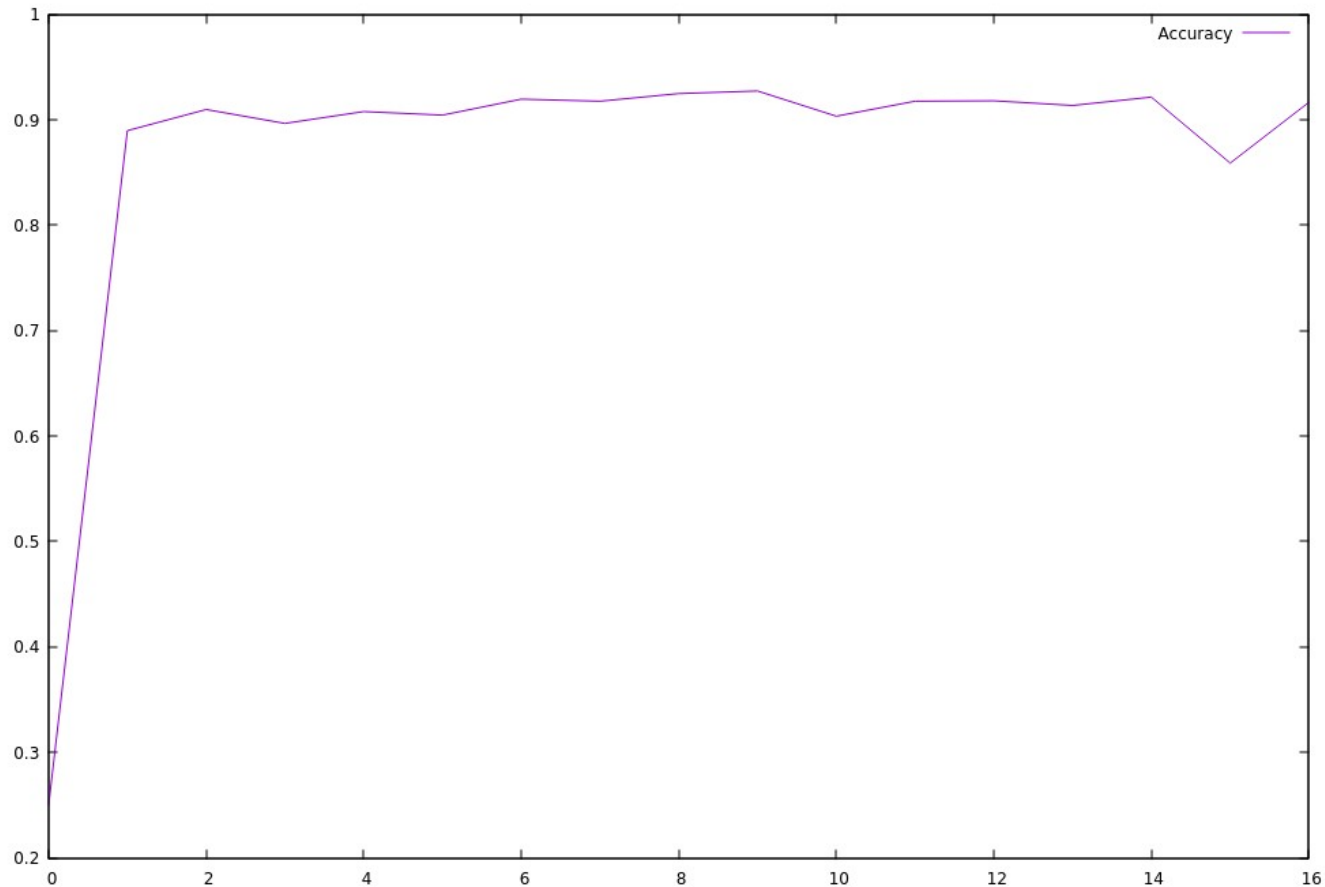
- » Pozwala prosto projektować warstwy sieci jako struktury dzięki wbudowanym funkcjom do przetwarzania tensorów
- » Posiada zaimplementowane funkcje optymalizujące

# Architektura sieci



# Trening

- » Trening odbywał się na CPU
- » Każda epoka trwała ok. 2h 45min



# Wynik treningu

| Kategoria | Brak kompresji | MP3 320kbps | MP3 192kbps | MP3 128kbps | Wszystko |
|-----------|----------------|-------------|-------------|-------------|----------|
| Zgodność  | 96,3%          | 93,4%       | 88,7%       | 88.3%       | 91,7%    |



# Zastosowanie uczenia maszynowego do wykrycia metody kompresji stratnej sygnału audio

<https://github.com/Ashymad/IOLA>

Dziękuję za uwagę

## Bibliografia:

- (1) Marina Bosi i Richard E. Goldberg: *Introduction to Digital Audio Coding and Standards*, Norwell, MA, USA: Kluwer Academic Publishers, 2002, isbn: 1402073577
- (2) Bongjun Kim: „Lossy Audio Compression Identification”, w: *2018 26th European Signal Processing Conference (EUSIPCO)*, 2018, s. 2459-2463
- (3) Romain Hennequin, Jimena Royo-Letelier i Manuel Moussallam: „Codec independent lossy audio compression detection”, w: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, s. 726–730