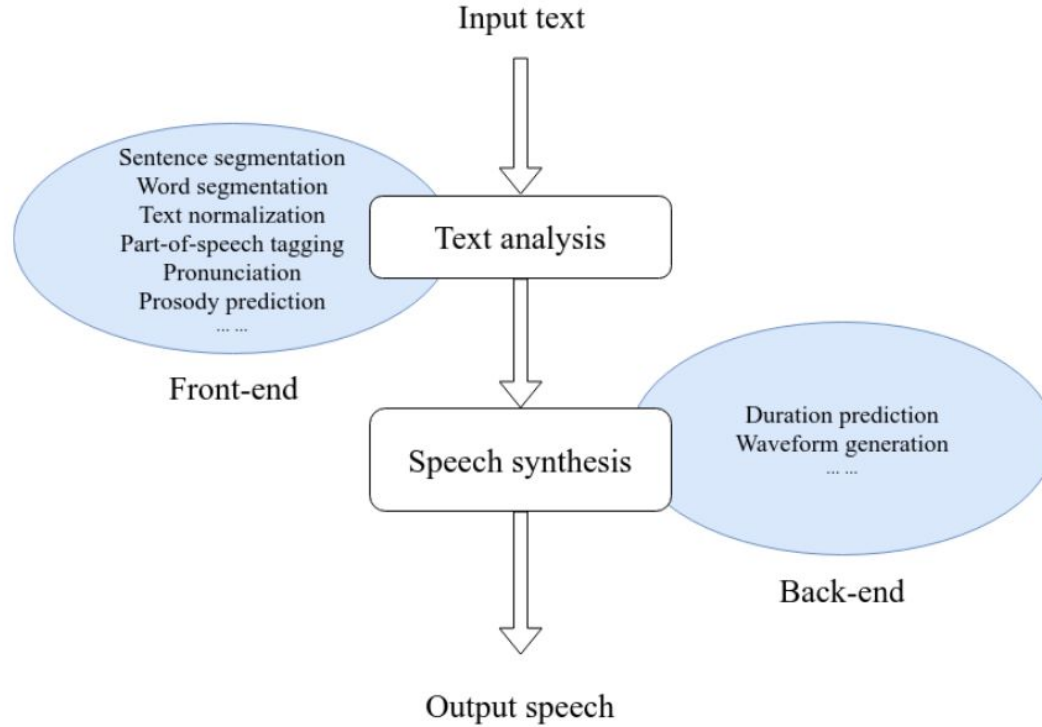


# Russian speech synthesis algorithms investigation based on spectrogram super resolution

scientific supervisor:  
Pavlovsky E. N.

student:  
Baushenko M. A.

# Introduction



# Recent Advantages

Sheng, L., Huang, D. Y., & Pavlovskiy, E. N. (2019). High-quality speech synthesis using super-resolution mel-spectrogram. arXiv preprint arXiv:1912.01167.

<b>Mel spectrogram</b>	<b>Griffin-Lim</b>	<b>WaveNet</b>
Coarse (baseline)	$3.29 \pm 0.069$	$3.84 \pm 0.061$
Predicted (ours)	$3.73 \pm 0.055$	$4.01 \pm 0.055$
Original	$3.74 \pm 0.054$	$4.04 \pm 0.056$

Note: ground truth:  $4.23 \pm 0.046$

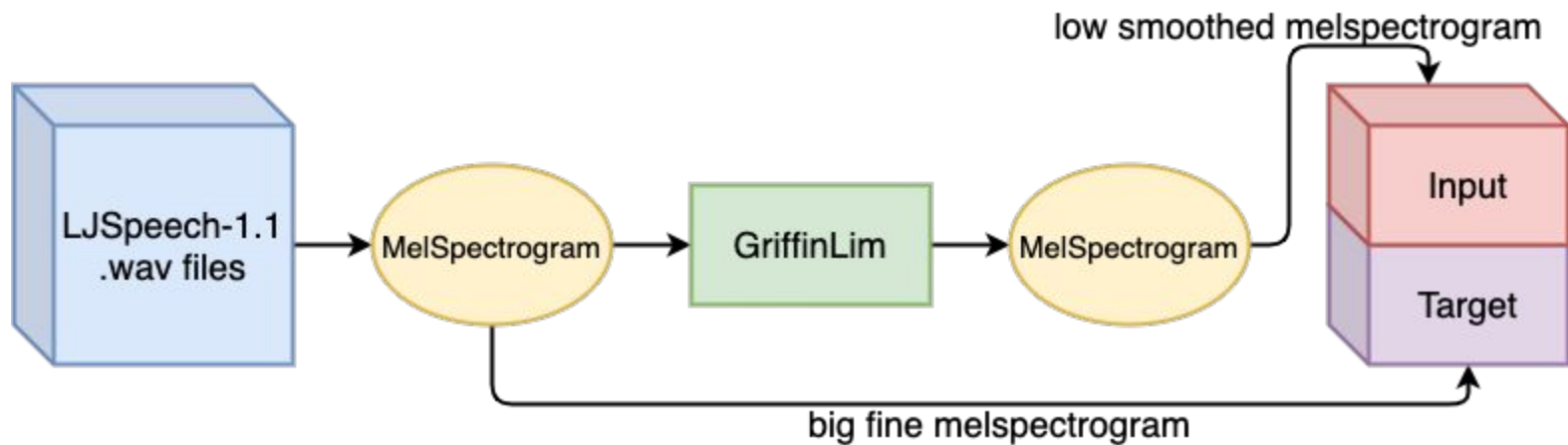
# Goals and objectives

Goal: improve the naturalness of speech by increasing the resolution mel-spectrograms

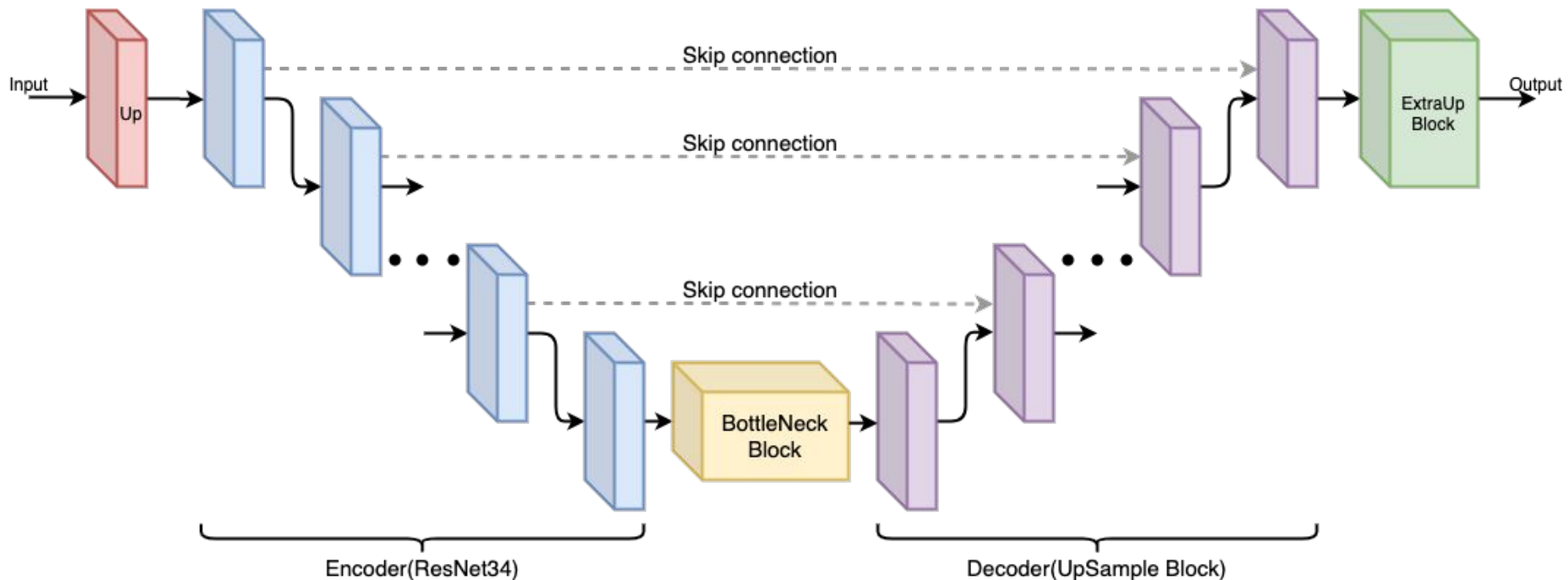
Objectives:

- Implement super-resolution preprocessing mel-spectrograms.
- Implement research pipeline.
- Train super-resolution preprocessing mel-spectrograms.
- Analyze the received metrics.

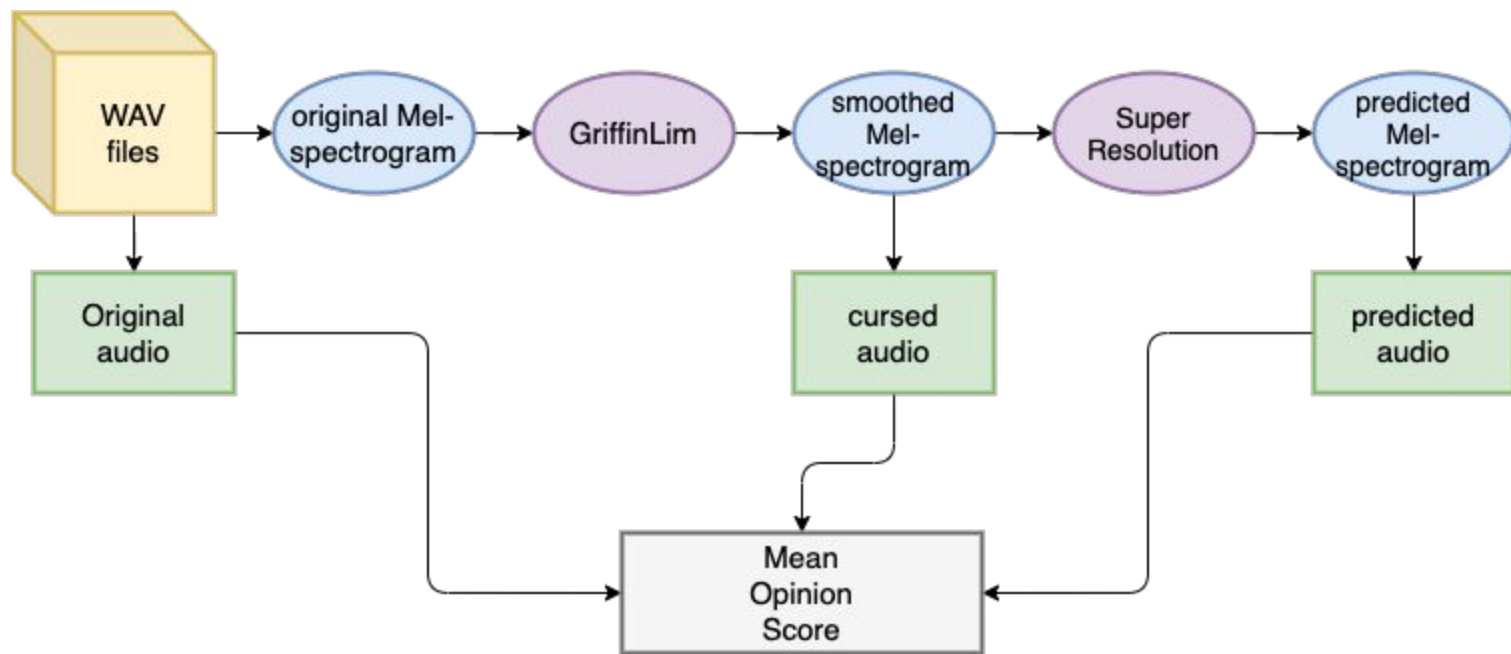
# Creating DataSet



# Architecture CNN U-Net



# Research pipeline



**Thanks for attention!**