

Voice Detection

1. Постановка задачи
2. Области применения
3. Метрики функции потерь
4. MarbleNet
5. CNN-Embedder + Self-attention.
6. Результаты.

Постановка задачи. Применение.

Задача выглядит следующим образом: у нас есть аудиозапись нужно определить или же сегментировать на участке где есть речь и где ее нету.

Применяется это вещь во многих местах:

1. Кодирование сигналов
2. Системы конференций.
3. Система распознавания речи.

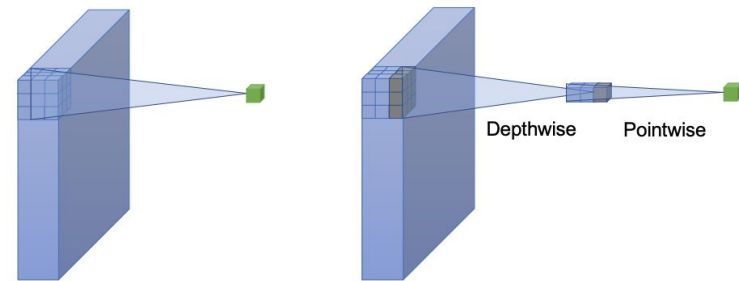
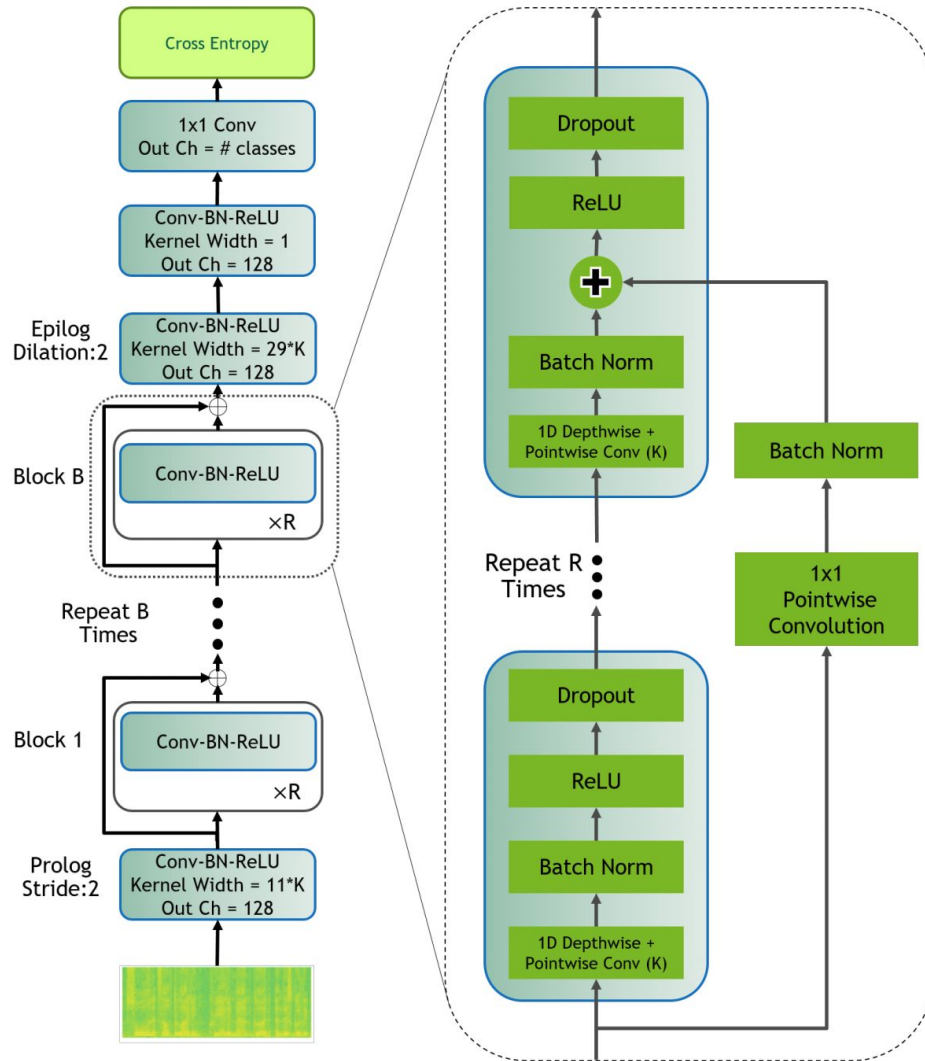
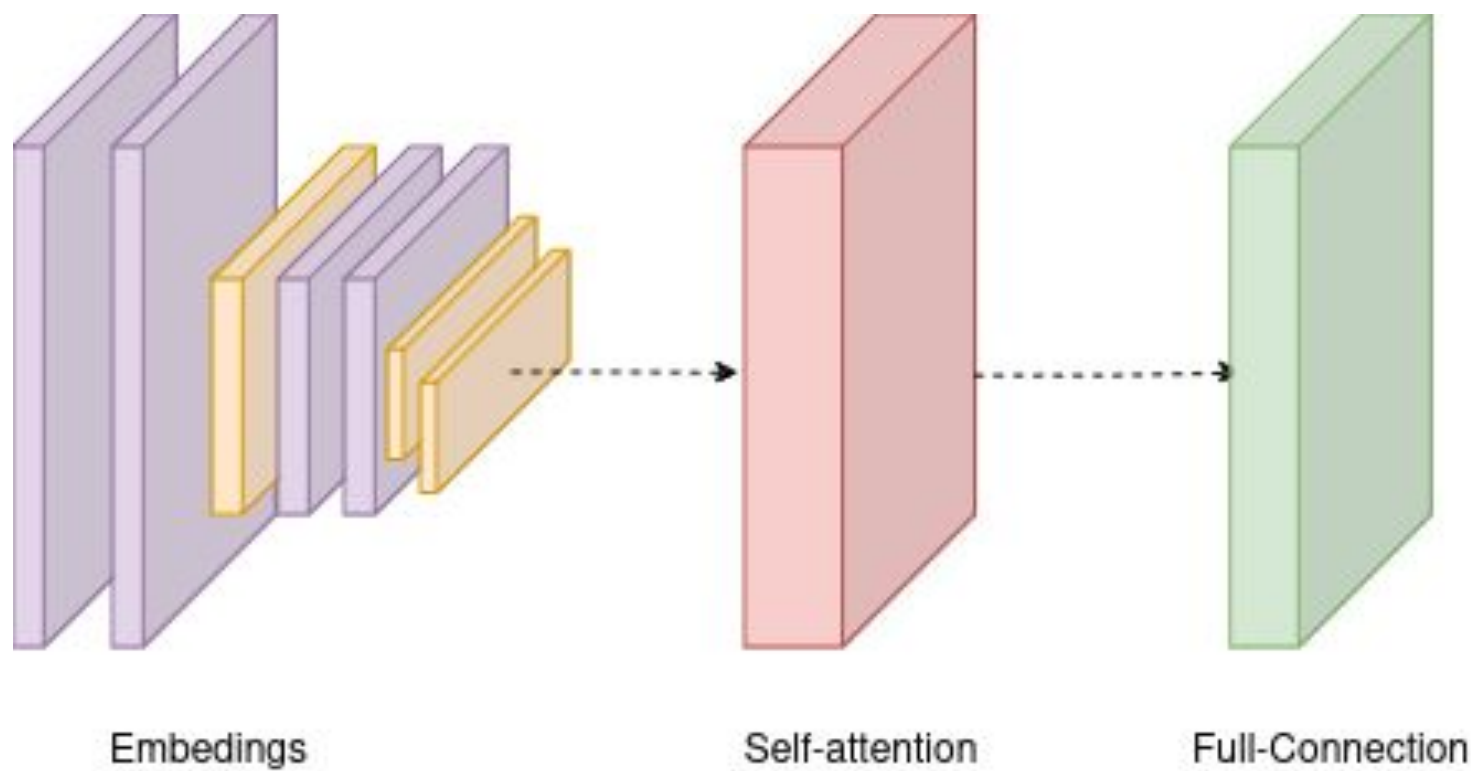
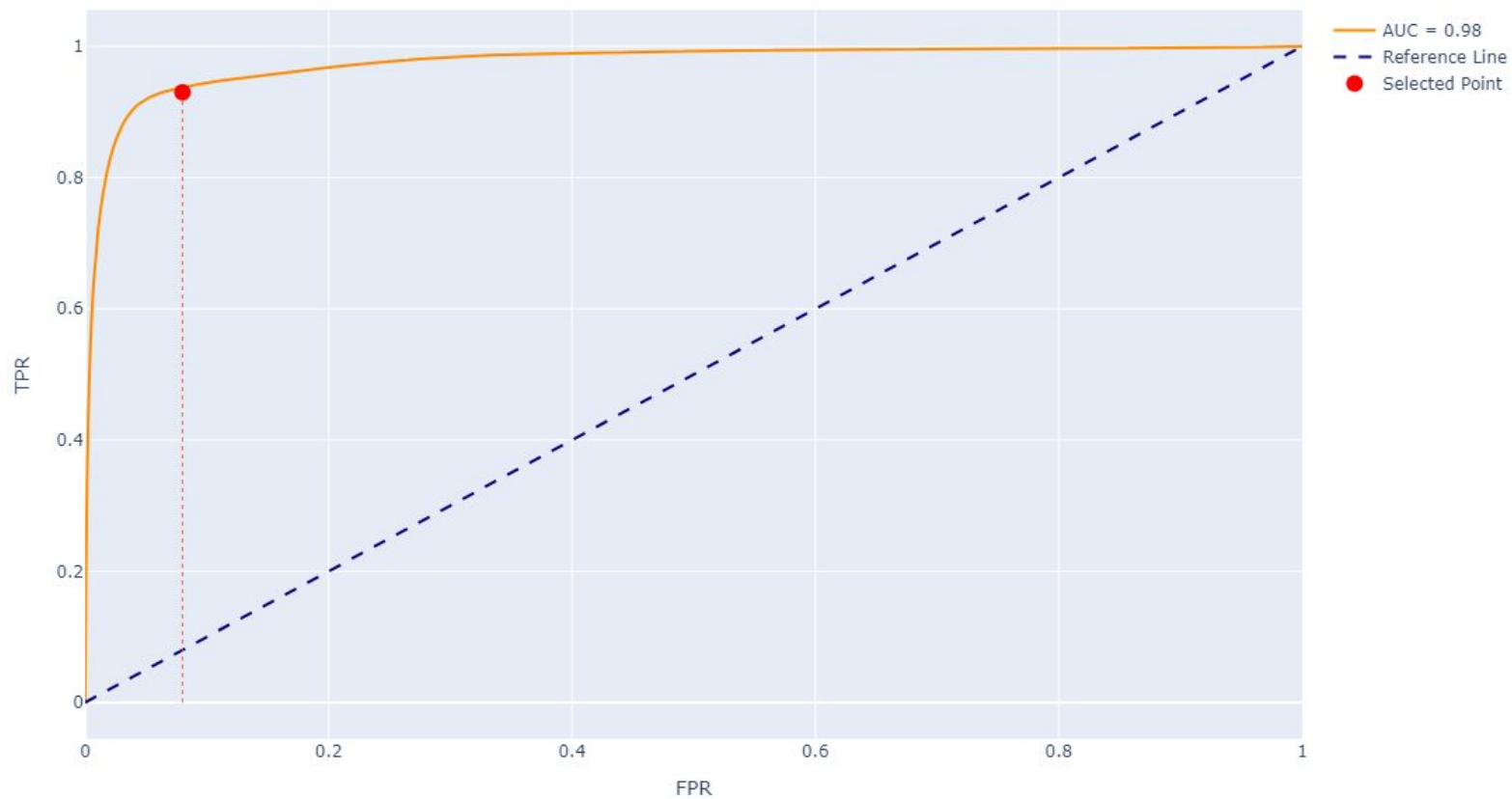
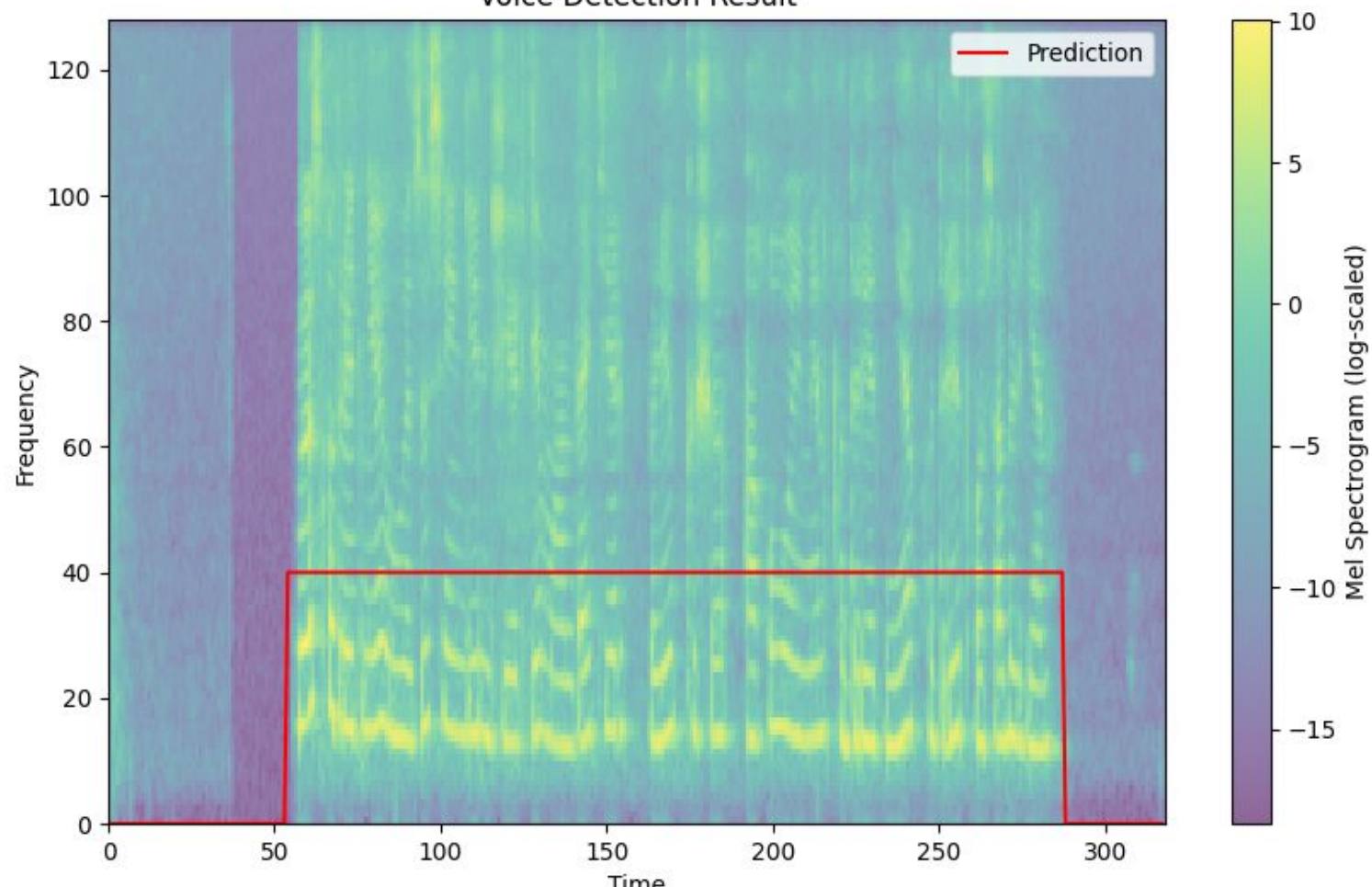


Figure 3: Standard convolution and depthwise separable convolution.





Voice Detection Result



Accuracy	Precision	Recall	F1 Score
0.9850	0.9915	0.9844	0.9829