Synthesizing Audio from Textual Input

Márcio Duarte
Luís Paulo Reis (Supervisor)
Gilberto Bernardes (Second Supervisor)

Faculdade de Engenharia da Universidade do Porto

Introduction
    Context
    Background
    Motivation
    Research objectives

- ▶ "Computer Science is the study of computation and information." [**university˙of˙york˙what˙nodate**]
- ▶ Evolution of Computer Science: From traditional programming to advanced Machine Learning (ML) and Deep Learning (DL) techniques.
- ▶ Importance of Big Data and Parallel Computing: Catalysts for advancements in ML and DL.
- ▶ Role of Deep Learning: Automated feature extraction, particularly effective in tasks like image and audio processing.
- ▶ Significance of Generative Models: Creating synthetic data for various applications.

1. **1958 - Birth of Modern AI:**
   - ► F. Rosenblatt proposes three fundamental questions leading to the development of the perceptron.
2. **1960s - Perceptron Convergence:**
   - ► Intensive work on convergence algorithms for the perceptron.
3. **1969 - Limitations of Perceptrons:**
   - ► Minksy and Papert demonstrate the limitations of perceptrons, leading to a slowdown in AI research.
4. **1980s - Emergence of Multilayer Neural Networks:**
   - ► Studies on learning under multilayer neural networks.
5. **1986 - Backpropagation:**
   - ► Rumelhart et al. describe backpropagation, a key learning procedure for neural networks.
6. **1990s - Second Winter of AI:**
   - ► Decreased investments in ML due to lack of real successes.
7. **Turn of the Millennium - Resurgence of ML:**
   - ► Emergence of three trends: Big Data, reduced cost of parallel computing, and interest in Deep Neural Networks (DNN).
8. **2010s - DL in Everyday Applications:**
   - ► DL becomes integral for various computer-made tasks,

1. **ML in Audio Processing**
   - ML techniques enhance sound synthesis, restoration, and speech recognition.
   - Learn complex patterns, improving quality and efficiency.
2. **Revolutionizing Sound Generation**
   - Integration of ML transforms sound creation and experience.
   - Opens creative avenues for artists, impacts industries like film, gaming, VR.
3. **Need for Further Research**
   - Urgency for studies in sound generation technologies.
   - This dissertation contributes significantly, offering resources for exploration.
4. **Impact and Contribution**
   - Reshaping human potential in sound creation through digital technologies.
   - Valuable resource for audio processing professionals, guiding future endeavors.
5. **Overall Significance**
   - Valuable contribution to audio processing and machine learning field.

1. Make a study of the current state-of-the-art deep learning architectures, focusing on generative ones.
2. Examine prior algorithms that can process sound for augmentation, feature extraction, or other purposes.
3. Make a study of the current state-of-the-art architectures used to develop sounds artificially.
4. Develop end-to-end systems that can synthesize sound from any given text input, while accounting for hardware constraints and ensuring reliable performance.
5. Evaluate the systems' ability to generate a sound from the given textual input accurately.

- ▶ Motivation
- ▶ Research question
- ▶ Contributions
- ▶ Outline