NARCISSUS Project

Surk Park

2018. 02.18

Agenda

- Speech Enhancement GAN
- NARCISSUS Project
- NARCISSUS Demo

Speech Enhancement GAN

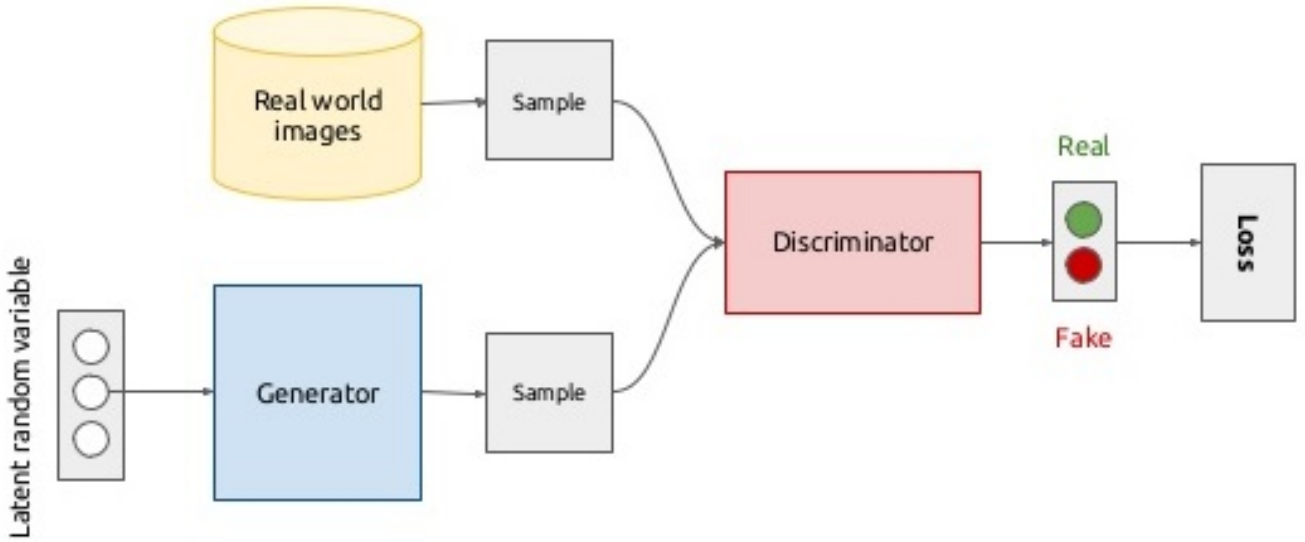
Speech Enhancement Techniques

Classic Methods and Current Deep Learning Methods for Speech Enhancement

- Classic Methods
  - ✓ Spectral subtraction
  - ✓ Wiener filtering
  - ✓ Statistical model-based
  - ✓ Subspace algorithm
- Deep Learning Methods
  - ✓ Denoising auto-encoder (RNN, LSTM, etc.)
    - Adding Noise feature into Input feature
    - Fourier analysis

Generative Adversarial Networks

GAN(Generative Adversarial Networks) Review

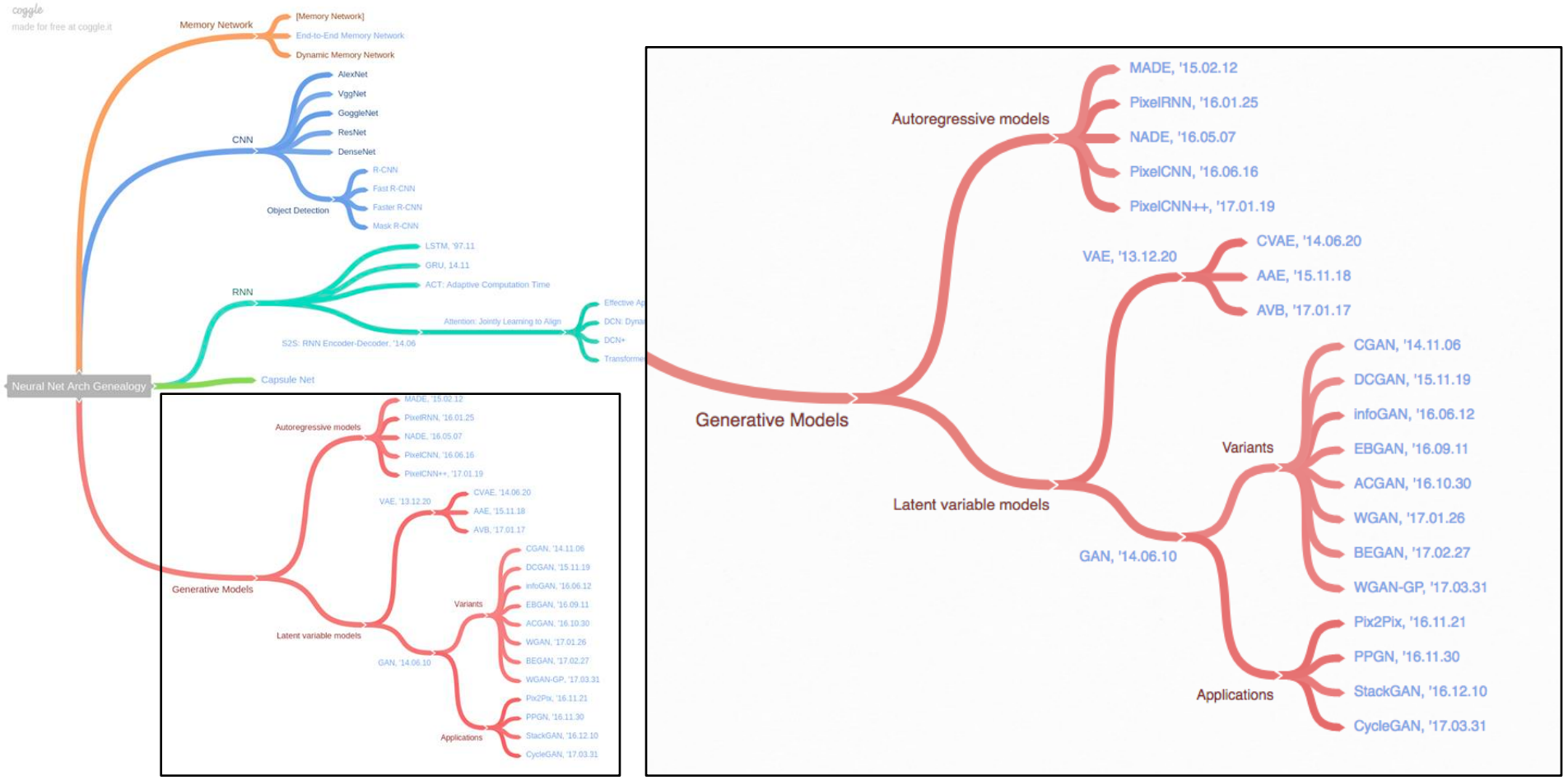


$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

- Generator: Fake sample Generate
- Discriminator: Fake sample Discriminate

Deep Learning Architecture Genealogy

Deep Leaning Architectrue Genealogy Review

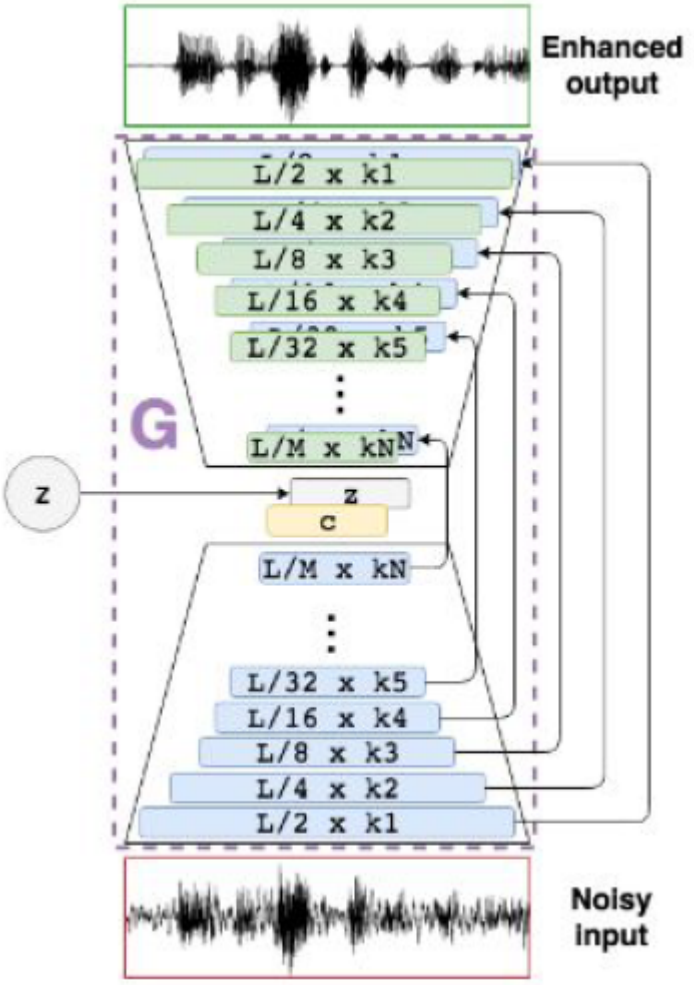


- Project page: https://github.com/hunkim/deep_architecture_genealogy
- Mind map https://coggle.it/diagram/Wf5mYoJbsgABUF9P

SEGAN : Speech Enhancement GAN

SEGAN(Speech Enhancement GAN) Introduction

- Using Raw data itself
- No Recursive operation like RNN
  - ✓ Fully-convolutional structure
- No definition like Noise-type



Original Paper Reference : https://arxiv.org/abs/1703.09452

SEGAN Training Result

Noise Cancellation Example

Speech Data Example	Raw Data	Speech Enhanced Data
(English, Woman) The orchestra was already increasing the scope of its ambitions		
(English, Man) There will be no repeat of that performance, that I can guarantee		
(Korean) Call Center		
(Korean) Dialogue in class lecture		

NARCISSUS Introduction

NARCISSUS ?

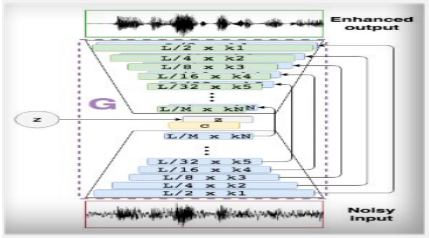
Smart Mirror Project applying Watson Cognitive APIs and SEGAN

NARCISSUS is a Smart-Mirror Project

- **uses Watson NLP/LT/STT/TTS AI APIs,**
- **converges other ML/DL based AI Services**
- **planned to make new interfaces with human beings and new MRR business models**

Speech Recognition

DL-based Speech Recognition + Noise Cancellation.



DL-based Speech Recognition

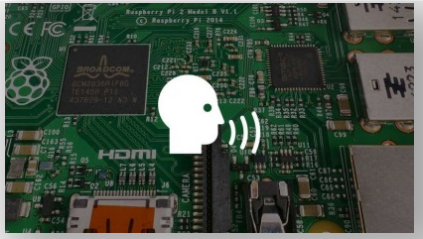
Watson STT API
Google Speech API

DL-based Noise Cancellation

Speech Enhancement GAN

Voice Control + Dialog

Interfaces with human by voice



Voice Control

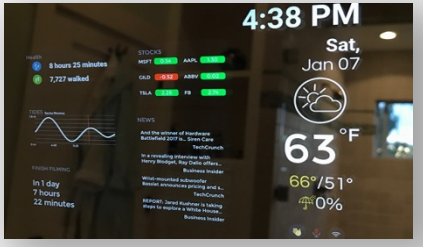
Open Source based Voice Control Module.

Dialog

Watson Conversation API.
Watson Natural Language Understanding API

Personal Assistant

Basic assistant function like Date/Time Reminder and External APIs connection like Map/Weather/News/SNS.



Basic Assistant Function

Data/Time Reminder.
UI Integration / Admin Console.

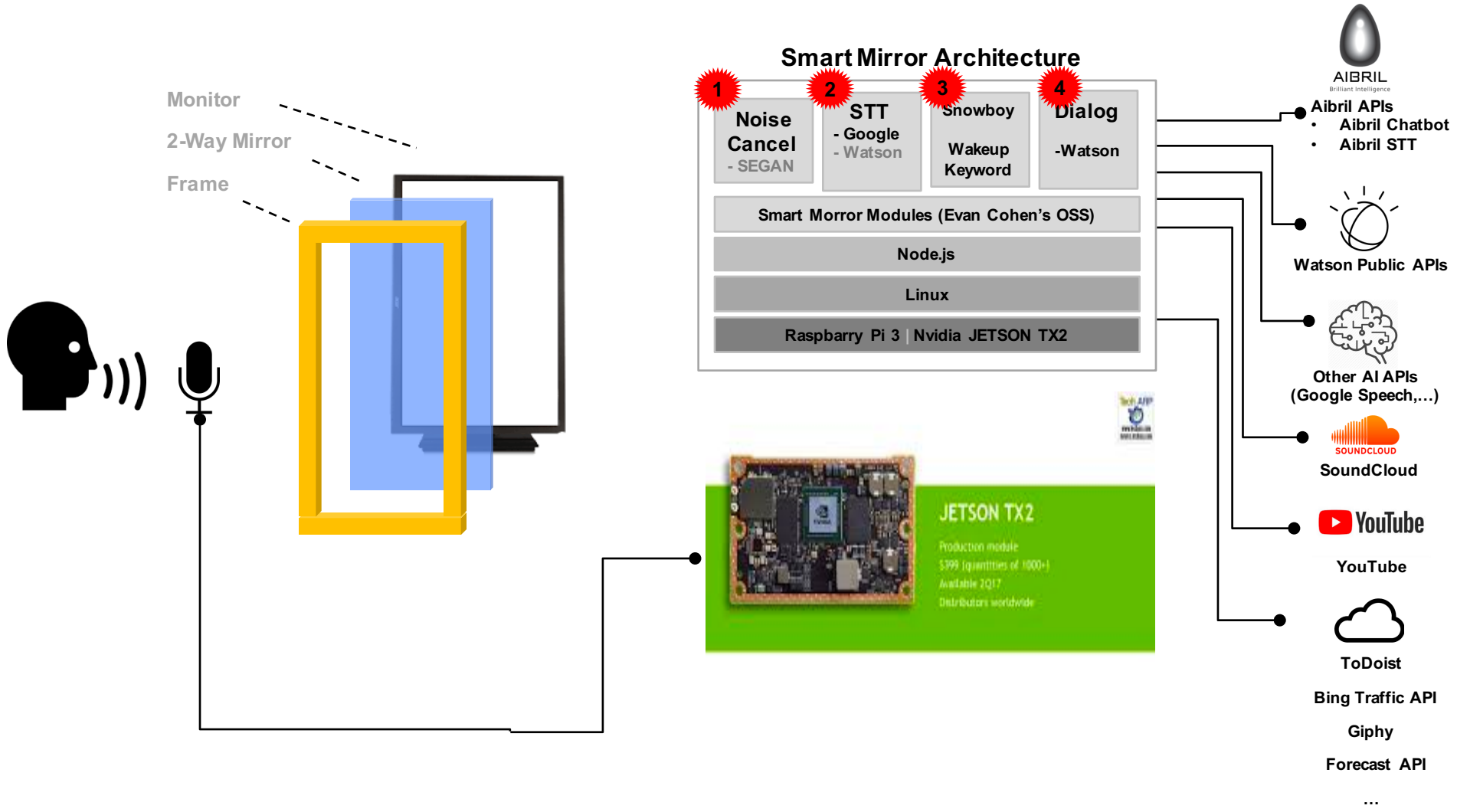
External APIs Connection

Map/Weather/News/SNS.
SoundCloud/Youtube …

Original Smart Mirror Reference : https://github.com/evancohen/smart-mirror

NARCISSUS Overall Diagram

DL-based Noise Cancellation + Voice Control + Dialog + External APIs + IoT



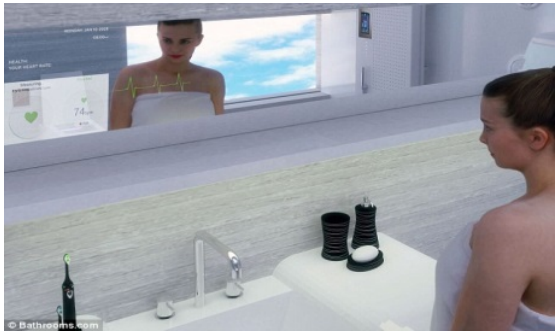
MRR BM by NARCISSUS

MRR Business Model by offering Cognitive APIs and Value-Added Services on NARCISSUS

Digital Concierge

Home Automation

Mirror Doctor



Digital Signage

Home Security

Home Mirror Cloud

NARCISSUS Demo

Demo

<NARCISSUS Project Demo 01 – Introduction>
https://youtu.be/hdTIafOrSow

<NARCISSUS Project Demo 02 – Voice Control>
https://youtu.be/YBw6hylnoYg

Thank you.