

The conference overview

Artem Sokolov
HSE, Nizhny Novgorod, Russia

SAMI 2019



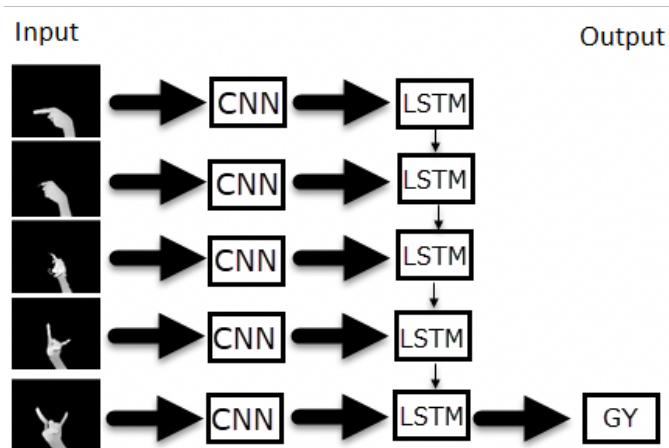
Herlany, Slovakia

Recognition of the Hungarian fingerspelling alphabet using Recurrent Neural Network

Венгерский язык жестов содержит 27 статических и 15 динамических жестов
Предыдущие решения использовали вариации сверточных сетей
Предлагается использовать рекуррентные слои для распознавания



Bence Dankó



Данные:

300 последовательностей по 10 кадров на каждую букву

Архитектура и результаты:

CNN+LSM+3FC -> softmax

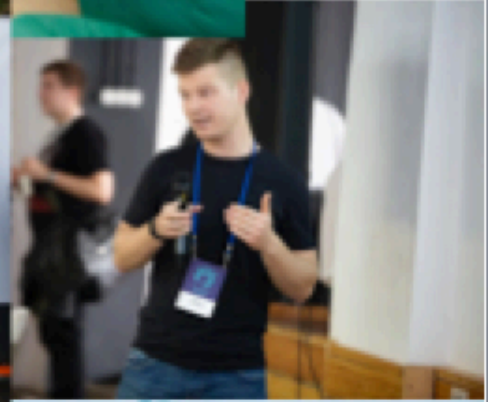
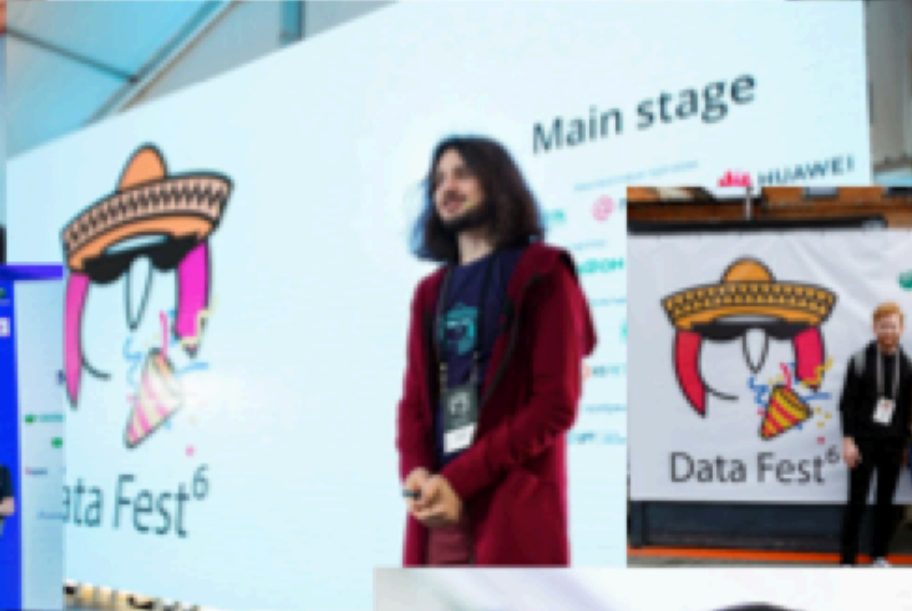
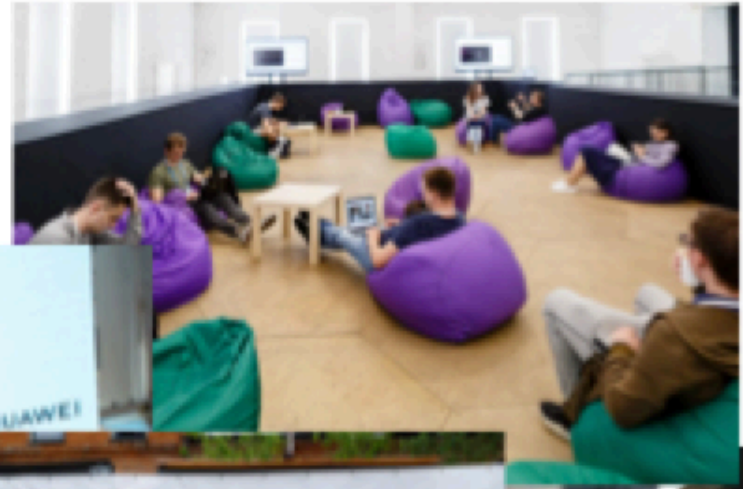
93,45% на тестовых данных

Datafest 2019

Moscow, Russia
May 10-11



DataFest



Робот-оператор в проде

Система автоматического обзвона. Напоминание о долге.

- Система должна уметь поддерживать некоторое количество сценариев, в том числе уметь переводить на оператора если что-то пошло не так.
- Голос должен быть максимально естественным
- Уметь отличать мужской голос от женского



Сергей Марков.
Сбербанк

Система:

- Kaldi (tdnn) + Tacotron 2 + Regex решение для понимания ответа должника + fasttest эмбедеднги для понимания некоторой окрестности слова
- Сверточная сеть для классификации голоса

Факты о подготовке:

- Сбербанк имеет в распоряжении ~миллиард записей разговоров с клиентами. Использовали 900 часов с покрытием разных регионов для тренировки акустической модели.
- Для генерации голоса на нормальном уровне достаточно 20 часов на английском. Но для русского много больше (сколько точно не сказано). Большая ручная работа проведена по подготовке intendoв и вычитыванию правильности грамматики и правописания (Потом тексты отдаются диктору.)
- Использование gan для того, что бы сделать речь более естественной

Thank you!