

# Методы глубинного обучения для анализа мультимодальных данных

Аспирант 1 года А. М. Грачев  
научный руководитель — к.т.н., доцент Д. И. Игнатов

ФКН ВШЭ  
Департамент анализа данных и искусственного интеллекта

23 июня 2016 г.

- Нейронные сети с большим количеством слоев
- Начали набирать популярность в середине 2000-ых годов
- Уже являются признанным индустриальным решением в нескольких областях: (обработка изображений, работа с аудио и другие)

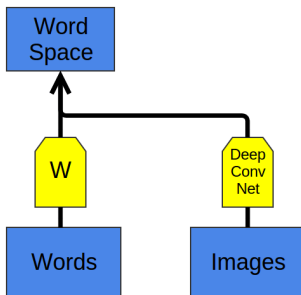
# Направления исследований

- 1 Работа над архитектурами
  - Интенсивный — усложнение базовых юнитов и внутренней структуры сети
  - Экстенсивный — расширение архитектуры сети
- 2 Развитие областей применимости
  - Новые задачи и структуры данных
  - Комбинирование модальностей

# Мультимодальные системы на базе методов deep learning

## Идея

Нейронная сеть может выучить представление объекта на изображении или слова в тексте. Давайте сматчим эти представления!



# Статьи по мультимодальные подходы



*Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, Andrew Y. Ng.* Multimodal Deep Learning — 2011



*Nitish Srivastava and Ruslan Salakhutdinov.* Learning Representations for Multimodal Data with Deep Belief Nets — 2012

# Пример NeuralTalk

Яркий пример — система NeuralTalk

<http://cs.stanford.edu/people/karpathy/neuraltalk/>



0.95	old
1.20	dump
2.36	truck
-0.01	sitting
-0.50	with
0.30	grass
-0.08	growing
-0.11	around
0.02	it

Ryan Pedone Photography

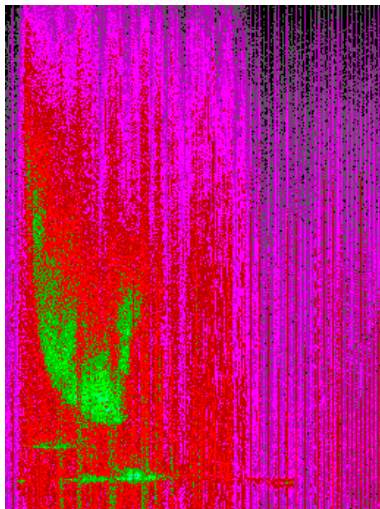
top sentences for this image:

- (4.14) An old dump truck sitting with grass growing around it.
- (3.57) A dump truck sitting by a barn surrounded by overgrowth.
- (3.43) A dump truck that is parked on some grass.
- (3.17) An old red train engine pulling cars behind it
- (3.02) A very old and rusted train parked on the tracks.

## Ионограммы. Постановка задачи

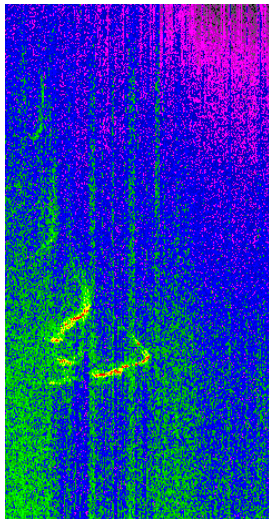
Ионограмма НЗИ состоит из развертки спектров сигнала с выхода ионозонда (т.е. является спектрограммой) и представляет собой зависимость амплитуды сигнала  $A$  от времени группового запаздывания  $\tau$  и частоты  $f$ . На ионограмме, кроме "полезного сигнала представленного треками различных мод распространения сигнала (протяженные по горизонтали наклонные кривые), присутствуют помехи различной природы и вида - шумовой фон ионограммы и вертикальные полосы (сосредоточенные помехи)

# Ионограммы. Пример

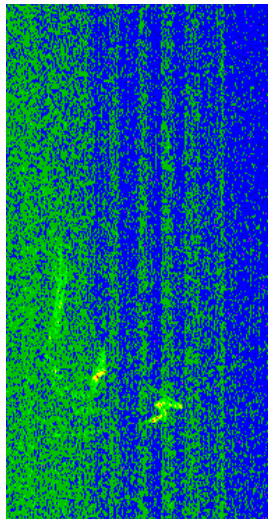




## Ионограммы. Еще один пример



a



b

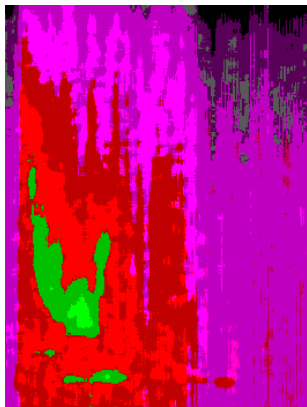
# Ионограммы. Проблемы

- Необходимость фильтрации
- Детектирование без учителя
- Классификация без учителя

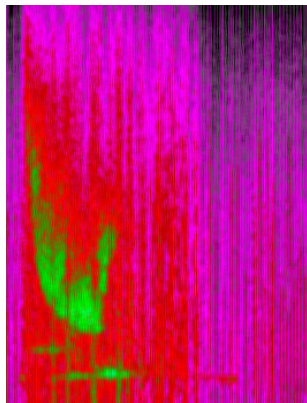
# Ионограммы. Схема решения

- 1 Простыми методами выделяем изображение — получаем слабую разметку
- 2 Обучаем нейронные сети, учимся выделять объекты
- 3 Кластеризуем автоэнкодерами на классы

# Ионограммы. Предобработка



a

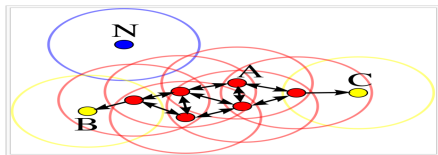


b

Рис.: a — медианный фильтр, b — фильтрация матрицей свертки

# Ионограммы. Методы

- 1 K-Means
- 2 DBscan
- 3 Mean shift



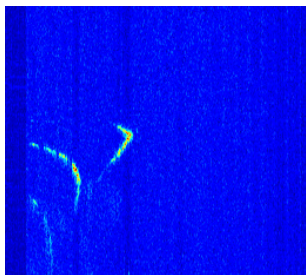
## Достоинства

- 1 Не надо задавать число кластеров
- 2 Может находить кластеры любой формы
- 3 Устойчив к выбросам
- 4 Требуется всего два параметра

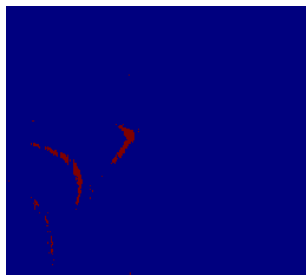
## Недостатки

- 1 Алгоритм недетерминированный
- 2 Требуется понимание масштаба и данных для задания расстояния  $\epsilon$  и числа точек  $N$
- 3

## Ионограммы. Результаты DBscan



a



b

Рис.: a — оригинальное изображение, b — результат применения dbscan'a

$$m(x) = \frac{\sum_{x_i \in N(x)} x_i K(x_i - x)}{\sum_{x_i \in N(x)} K(x_i - x)}$$

## Достоинства

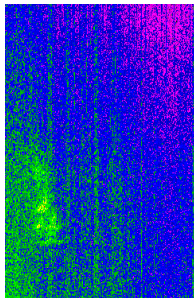
- 1 Не надо задавать число кластеров
- 2 Может находить кластеры любой формы
- 3 Зависит только от одного параметра: размера окна

## Недостатки

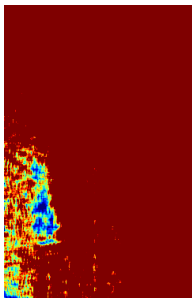
- 1 Выбор размера окна нетривиален
- 2 Часто требует адаптивного подбора размера окна



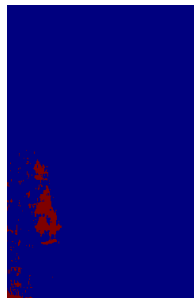
# Ионограммы. Результаты MeanShift



a



b










c

## Дальнейшая работа

- 1 Обучение нейронных сетей на базе вот такой слабой разметки
- 2 Теоретические исследования архитектуры

## Основная литература

-  *Li Deng and Dong Yu.* Deep Learning: Methods and Application — Foundations and Trends in Signal Processing Vol. 7, Nos. 3-4 (2013), Pages 197-387
-  *J. Schmidhuber.* Deep Learning in Neural Networks: An Overview — Neural Networks, Volume 61, January 2015, Pages 85-117
-  *Y. Bengio.* Learning Deep Architectures for AI — Foundations and Trends in Machine Learning 2009
-  *Y. Bengio and I. J. Goodfellow and A. Courville* Deep learning — Unpublished (book in preparation for MIT Press), <http://www.iro.umontreal.ca/bengioy/dlbook> 2015
-  *Y. LeCun, Y. Bengio and G. Hinton* Deep Learning — Nature, Vol. 521, pp 436-444., 2015
-  *A. Babenko, A. Slesarev, A. Chigorin, V. Lempitsky* Neural Codes for Image Retrieval — 2014
-  *A. Krizhevsky, I. Sutskever and G. E. Hinton* ImageNet Classification with Deep Convolutional Neural Networks — Advances in Neural Information Processing Systems 25 1090–1098, 2012