



NATIONAL RESEARCH
UNIVERSITY

КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ, ОПИСЫВАЕМЫХ НЕЙРОСЕТЕВЫМИ ПРИЗНАКАМИ ВЫСОКОЙ РАЗМЕРНОСТИ: ПОВЫШЕНИЕ ВЫЧИСЛИТЕЛЬНОЙ ЭФФЕКТИВНОСТИ БЕЗ СНИЖЕНИЯ ТОЧНОСТИ

Савченко А.В.

Доктор техн. наук, в.н.с. ЛАТАС, проф.
НИУ ВШЭ - Нижний Новгород

URL: www.hse.ru/en/staff/avsavchenko

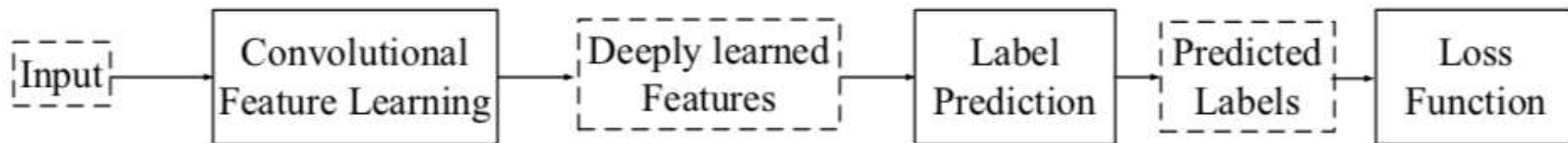
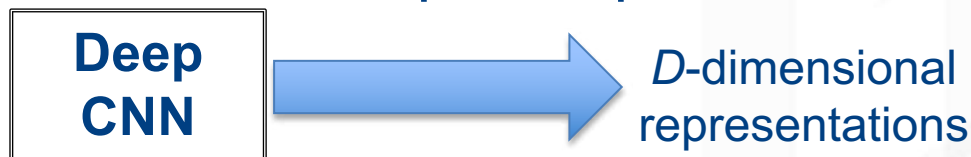
Международная конференция
«Нейронные сети послезавтра: проблемы и перспективы»
01.12.2019



Вычислительно эффективные методы распознавания изображений

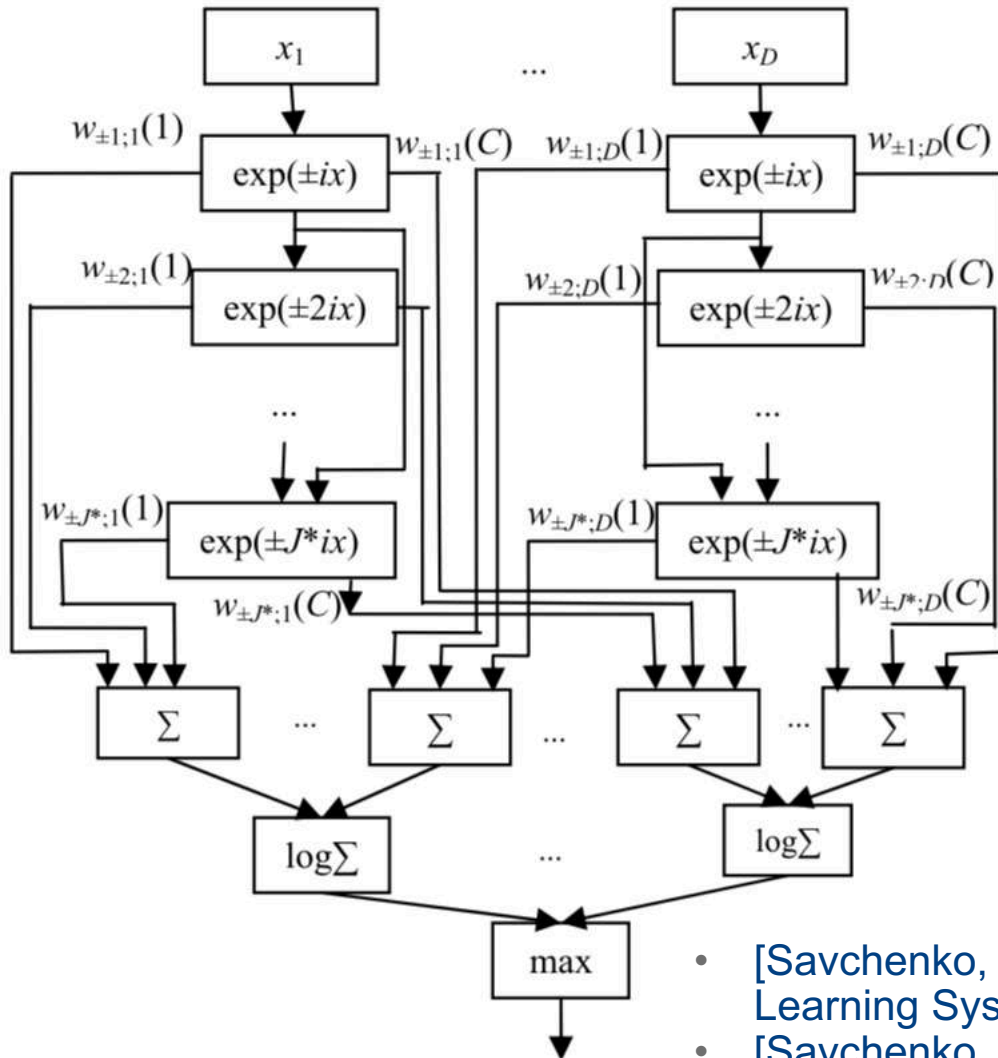
Задача распознавания изображений: $C \gg 1$ классов, R обучающих примеров (эталонов).

Число изображений каждого класса недостаточно для обучения сложного классификатора: $R/C < const$



Полный перебор (instance-based learning), сложность: $O(RD)$.

- Низкая вычислительная эффективность
 - Многокритериальная оптимизация (минимизируем вероятность ошибочной классификации и время принятия решений):
- $$\bar{\alpha} \rightarrow \min \quad \bar{t} \leq t_0$$

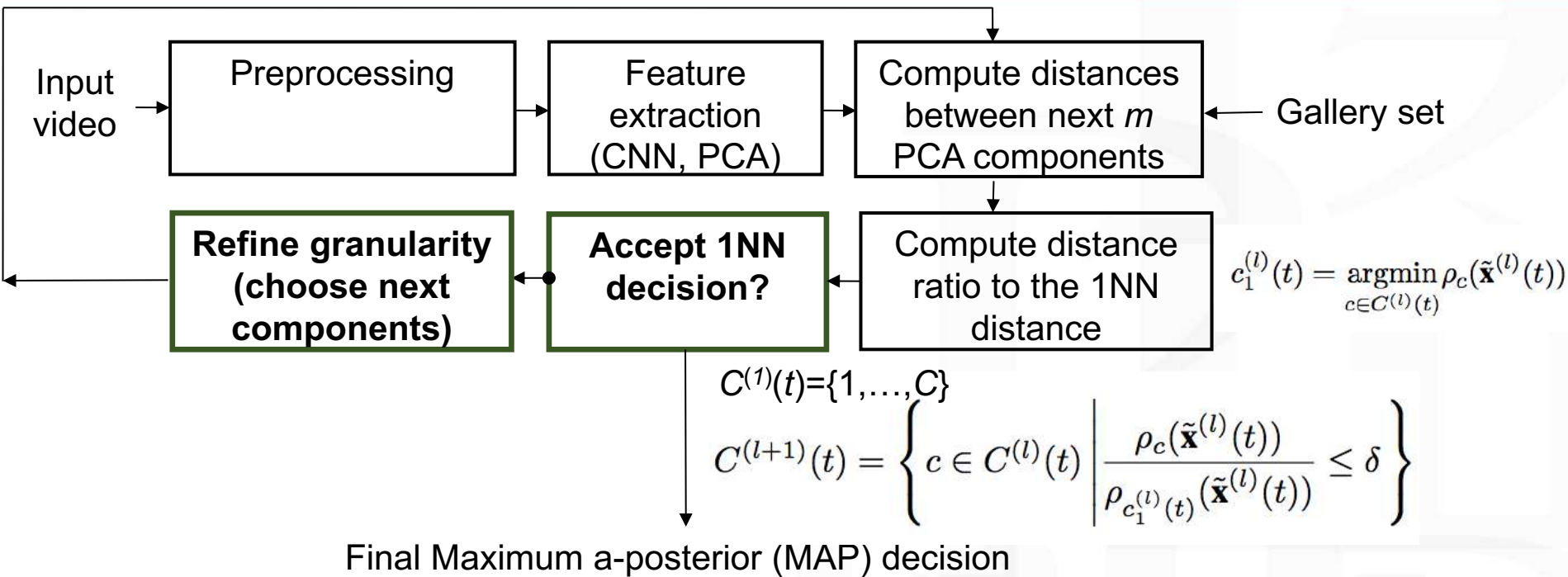


- Сходится к Байесовскому решению.
- Высокая скорость обучения: новый обучающий пример добавляется в реальном времени
- Классификация **точнее и примерно в $(R/C)^{2/3}$ -раз быстрее**, чем instance-based learning (PNN, k-NN)

Проблемы

- Медленнее PNN при $R/C < 5$
- Менее точна, чем SVM, RF, ... при $R/C > 25$

- [Savchenko, IEEE Transactions on Neural Networks and Learning Systems, 2019]
- [Savchenko, ICPR 2018]



- В разы быстрее полного перебора
- Значимо точнее, чем сравнение главных компонент

Проблема

- Точность на 0.1-0.5% **меньше** PNN/k-NN
 - [Savchenko, Information Systems, 2019]
 - [Savchenko, Knowledge-Based Systems, 2016]

Метод максимального правдоподобного перебора

$$r_{k+1} = \underset{\nu \in \{1, \dots, R\} - \{r_1, \dots, r_k\}}{\operatorname{argmax}} \left(p_\nu \cdot \prod_{i=1}^k f(\rho(X, X_{r_i}) | W_\nu) \right),$$

Дивергенция Кульбака-Лейблера асимптотически распределена как нецентральный хи-квадрат.

$$r_{k+1} = \underset{\mu \in \{1, \dots, R\} - \{r_1, \dots, r_k\}}{\operatorname{argmin}} \left(\sum_{i=1}^k \varphi_\mu(r_i) - \ln p_\mu \right) \quad \varphi_\mu(r_i) \approx \frac{(\rho(X, X_{r_i}) - \rho_{\mu, r_i})^2}{\rho_{\mu, r_i}}$$

- На порядок быстрее k-NN/PNN

Проблемы

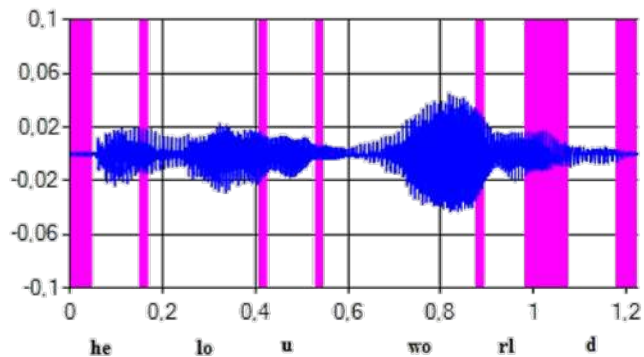
- Точность по сравнению с полным перебором снижается на 0.2-1%
- Дополнительные затраты памяти: не масштабируется на сотни тысяч классов
- [Savchenko, Pattern Recognition, 2017]
- [Savchenko, Optimization Letters, 2017]
- [Savchenko, Pattern Recognition, 2012]



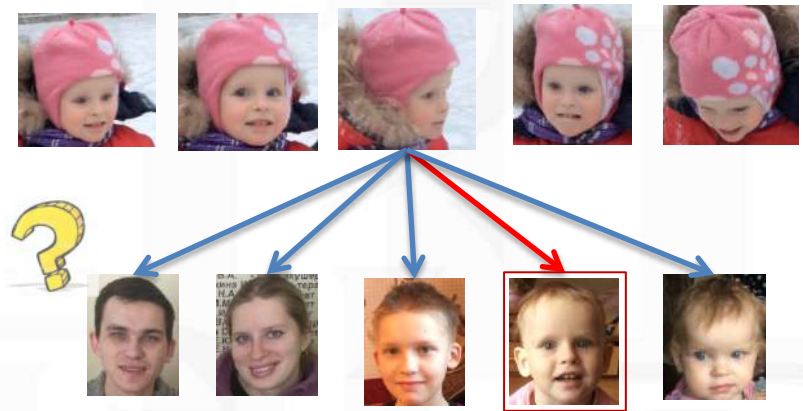
Классификация набора (последовательности) объектов

На вход классификатора подается последовательность $\{X(t)\}$ из $T > 1$ наблюдений одного объекта

Распознавание речи



Распознавание видео



- Рекомендации по фото купленных/просмотренных товаров
- Распознавание событий в галерее фотографий на мобильном устройстве
- ...

$$f(\rho_1(\mathbf{x}), \dots, \rho_C(\mathbf{x}) | W_c) = f(\rho_c(\mathbf{x}) | W_c) \cdot \prod_{i=1, i \neq c}^C f(\rho_i(\mathbf{x}) | W_c)$$

Добавление регуляризации к методу ближайшего соседа

$$\min_{c \in \{c_1, \dots, c_M\}} \left(\rho_c(\mathbf{x}) + \frac{\lambda}{C} \sum_{i=1, i \neq c}^C \frac{(\rho_i(\mathbf{x}) - \rho_{c;i} - \frac{D-1}{WH})^2}{\rho_{c;i} + \frac{D-1}{4WH}} \right)$$

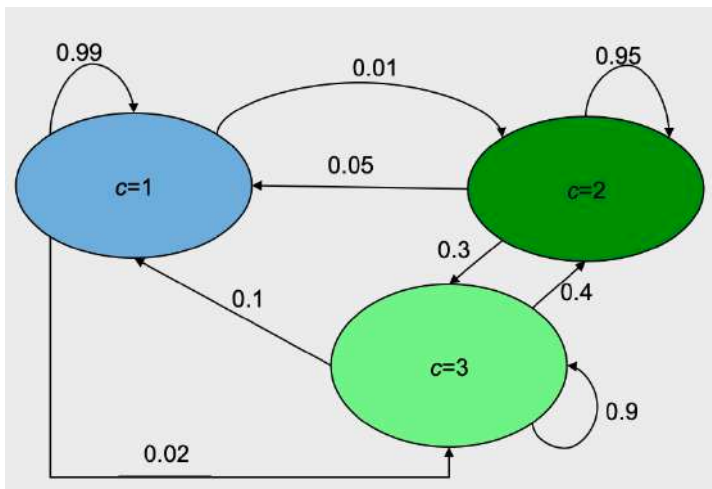
- Значимое повышение точности для классификации набора изображений

Проблема

- Дополнительные вычислительные затраты по сравнению с k-NN

[Savchenko, Belova, Expert Systems with Applications 2018]

Каждый класс и распознаваемая последовательность представляются как нечеткие множества эталонных обучающих примеров



$$\mu(r, t) = \mu_{\mathcal{V}(t)}(\mathbf{x}_r^*) \mu(\mathbf{x}_r^* | \mathbf{x}(t)).$$

$$\mu_i(\mathbf{x}_r^*) = \frac{\exp(-\lambda \rho_{KL}(\mathbf{x}_i^*, \mathbf{x}_r^*)) p_r}{\sum_{k=1}^R \exp(-\lambda \rho_{KL}(\mathbf{x}_i^*, \mathbf{x}_k^*)) p_k}$$

$$\mu(\mathbf{x}_r^* | \mathbf{x}(t)) \stackrel{def}{=} \hat{P}(\mathbf{x}_r^* | \mathbf{x}(t)) = \frac{\exp(-\lambda \rho_{KL}(\mathbf{x}(t), \mathbf{x}_r^*)) p_r}{\sum_{k=1}^R \exp(-\lambda \rho_{KL}(\mathbf{x}(t), \mathbf{x}_k^*)) p_k}$$

- Значимое повышение точности для классификации фонем

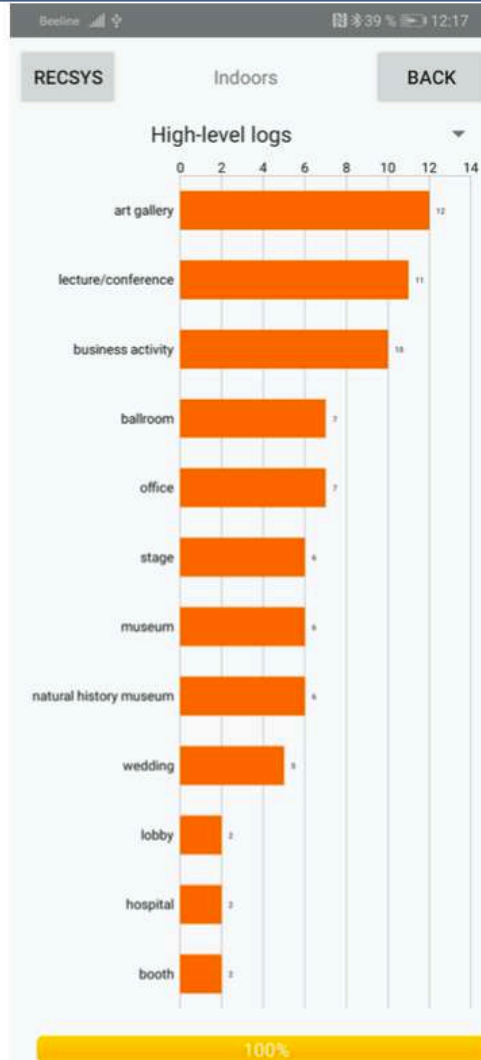
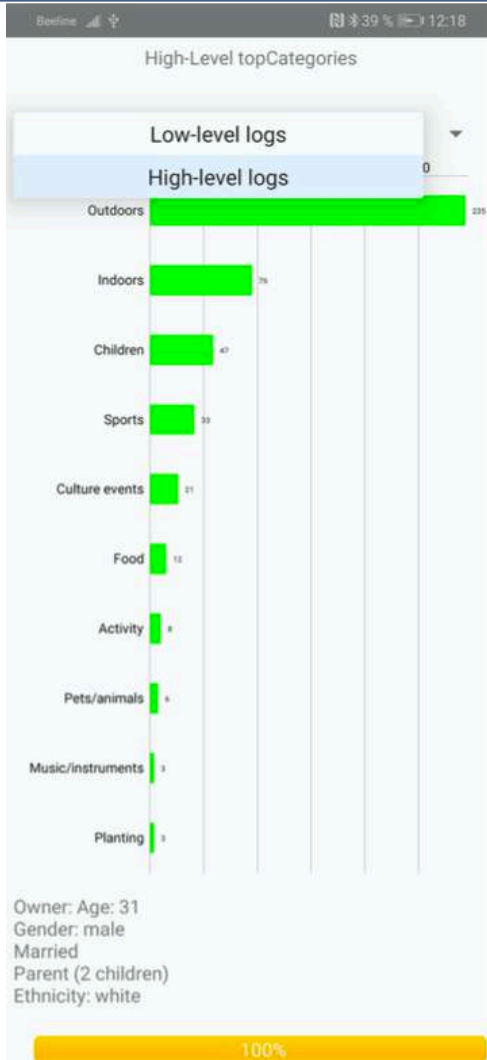
Проблема

- Дополнительные вычислительные затраты по сравнению с k-NN

- [Savchenko A.V., Savchenko L.V. Pattern Recognition Letters, 2015]
- [Savchenko A.V. et al. Optical Memory and Neural Networks, 2018]



Извлечение предпочтений пользователя мобильного устройства



Sep 8, 2018

Sep 25, 2019

Sep 11, 2018

Jun 27, 2018

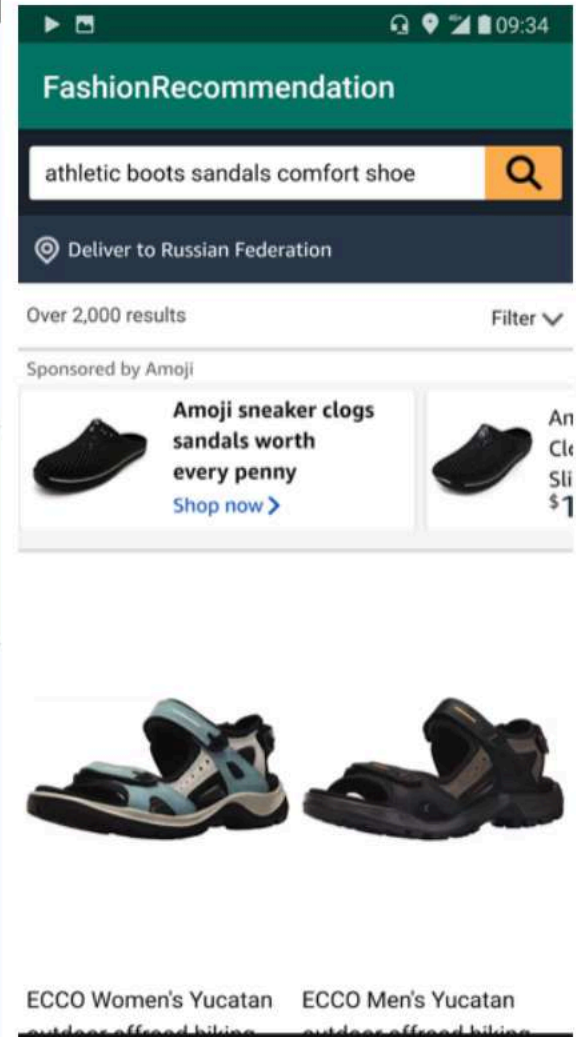
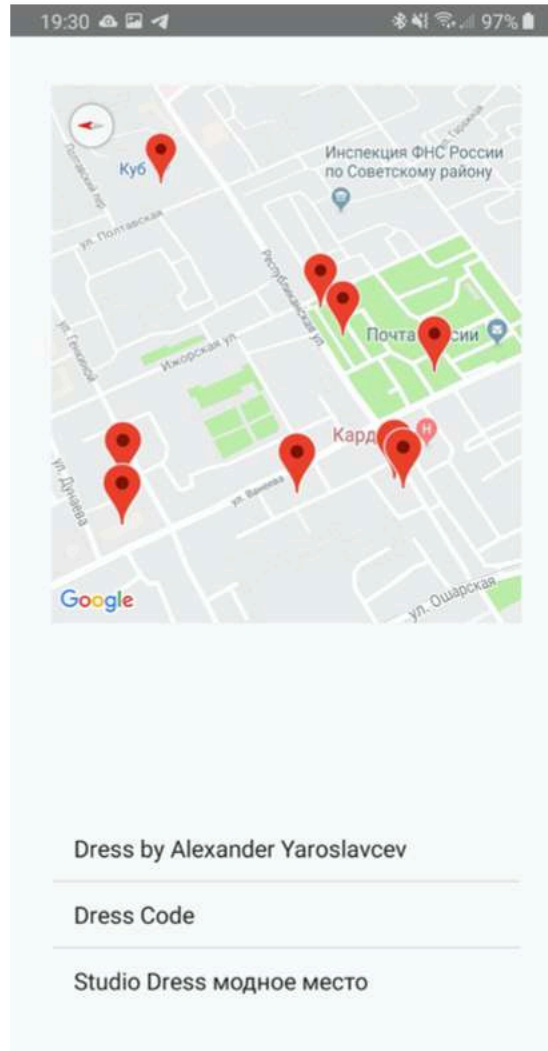
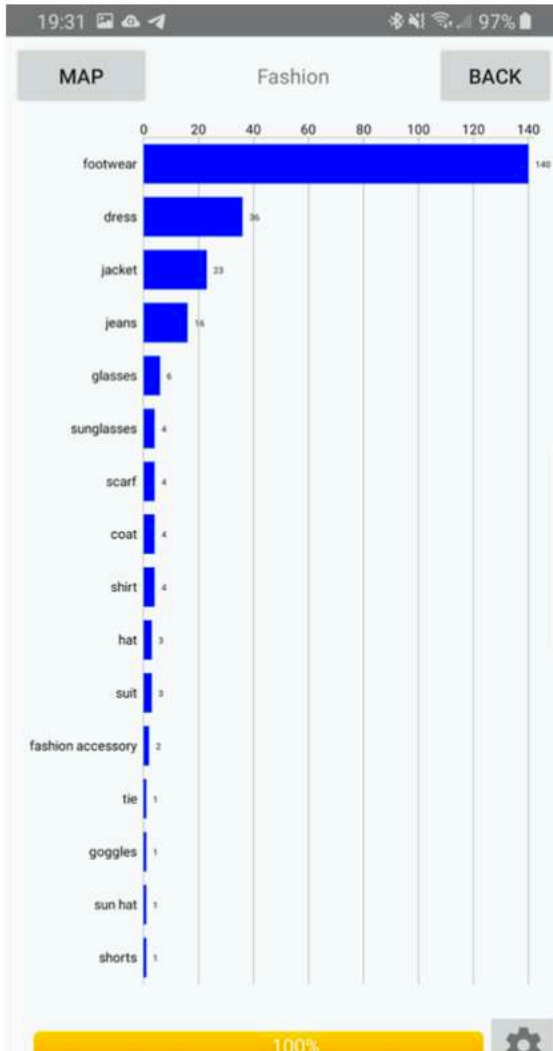
Aug 31, 2019

PREV NEXT BACK

photo 1 out of 2
Public photo
latitude=0,000 longitude=0,000

no objects found
scenes: football (0,98); events: sports (0,17);
No faces found
text:OAMMn JAKO
Instat
!!
OOAMH OL
BAdSIRRU
1AAIH
OLMP.BET

100%



- 1. Типично повышение вычислительной эффективности за счет некоторого снижения точности**
- 2. Возможны сценарии, в которых при дополнительных ограничениях на обучающее множество, повышается как точность, так и быстродействие.**
- 3. Характерные признаки извлекаются в глубоких нейронных сетях. Время прямого прохода (inference) зачастую намного превышает время классификации.**



NATIONAL RESEARCH
UNIVERSITY

Спасибо за внимание!