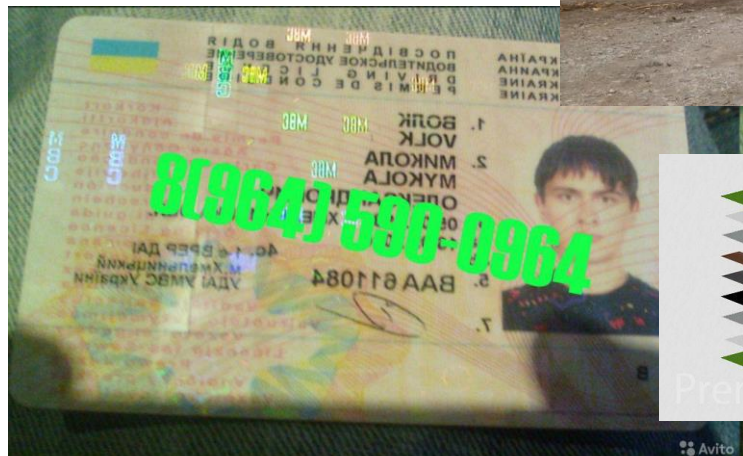
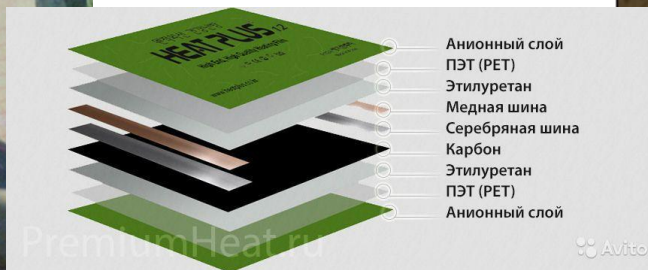


Конкурс по распознаванию контактной информации на изображениях [Avito.ru-2014](#)

# Подход к решению задачи

Призер конкурса Ульянов Дмитрий

# Примеры изображений

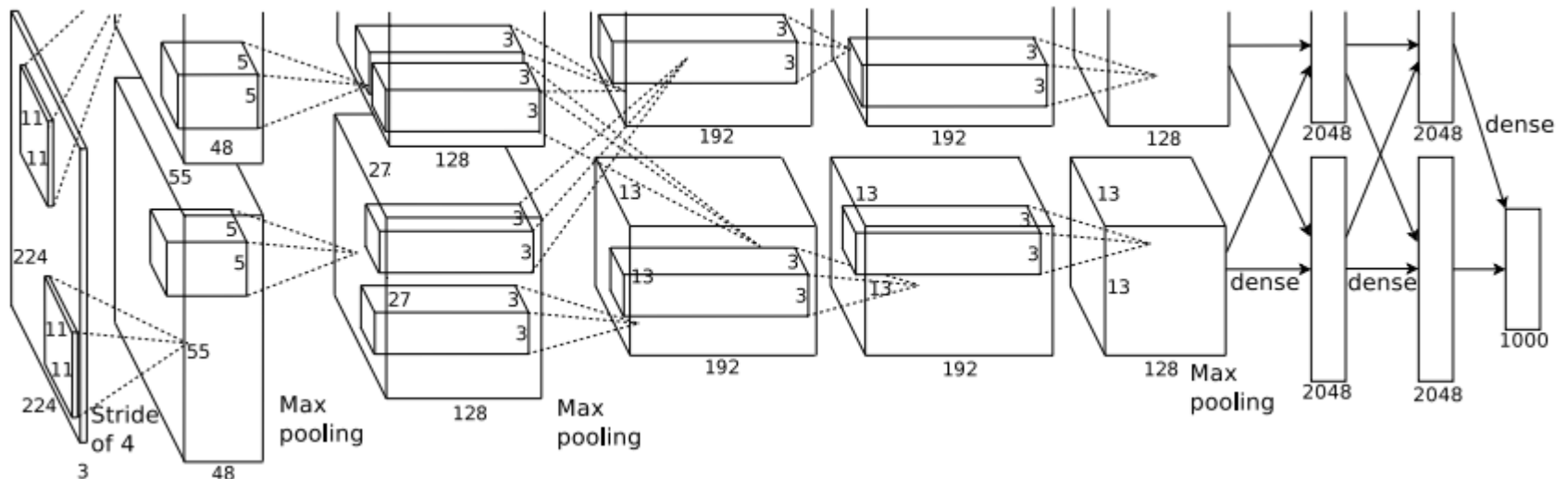


**ВСЕ**  
**ПЕРЕВОЗКИ**  
**НЕ ДОРОГО**  
**8 (964) 713-92-92**



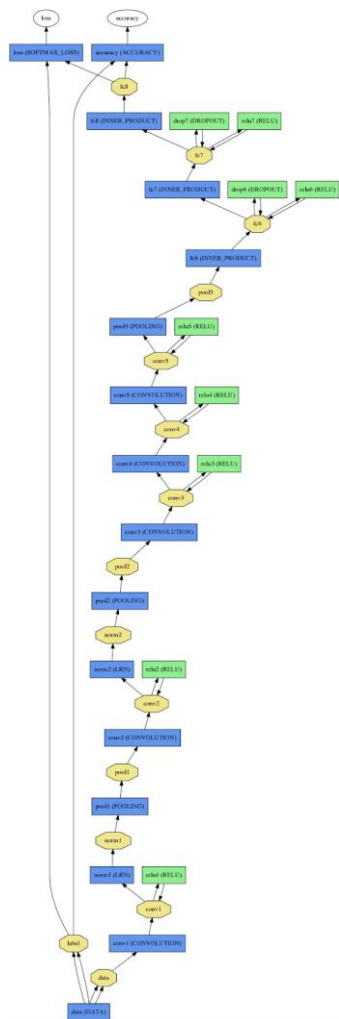
# Глубинное обучение, AlexNet

- В 2012 году А. Krizhevsky обучил нейросеть **AlexNet**, которая выиграла конкурс по классификации изображений базы *ImageNet* на 1000 классов.
- Структура сети:
  - Сверточные слои + max pooling + локальная нормализация
  - Полносвязные слои в конце + dropout
  - ReLU в качестве функций активации
  - Обучение длилось неделю.
  - Более подробно [тут](#).



# Глубина обучения

- За два года нейросети стали еще глубже (представлено качество сетей на базе ILSVRC-2012)



← *AlexNet* (2012) 18.2% err

*GoogLeNet* (2014) 10.7% err →



# Извлечение признаков

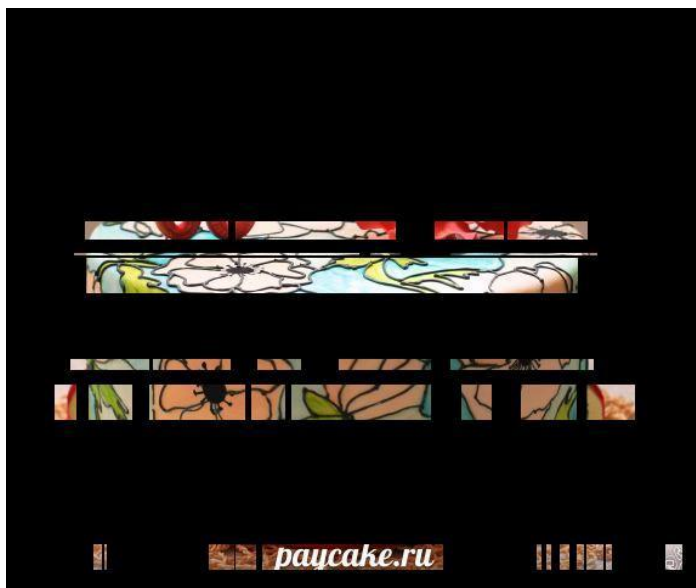
- Основная идея глубинного обучения (Deep Learning) -- автоматическое извлечение сложных признаков
  - В рамках нейросети одновременно извлекаются признаки и настраивается классификатор, что лучше, чем делать эти два этапа отдельно.
- Многие задачи решаются по принципу
  - Взять обученную по базе ImageNet модель
  - Переобучить последний слой на анализируемых данных с фиксированными другими слоями
  - Дообучить всю сеть (finetuning) уменьшив шаг.
- Использовать активации с некоторого слоя сети как признаки для анализируемых изображений
  - Прогоняем данные через нейросеть, получаем новую выборку с признаками.
  - Решаем обычную задачу классификации.

# Изображения с маской

- Искалась возможность помочь классификатору найти связь между содержимым изображений и метками.
- Контактная информация – это текст. Поэтому можно попробовать выделить текст на изображениях.
- Настроим детектор так, чтобы он не пропускал текст, но, возможно, выделял что-то лишнее.



# Изображения с маской



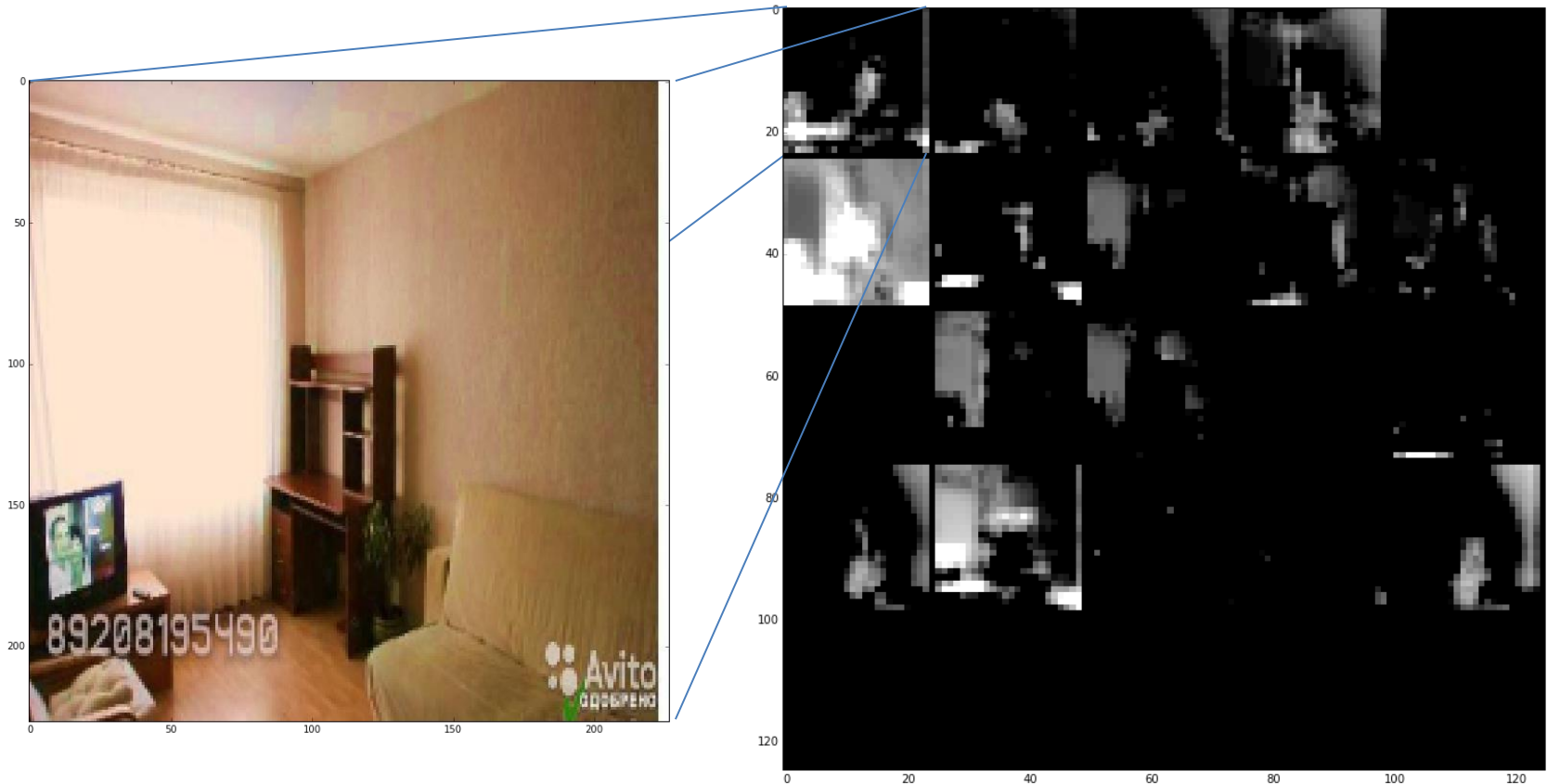
# Решение задачи конкурса

- Признаки
  - С первого полносвязного слоя общедоступной ImageNet сети
  - + 900 признаков HOG от изображений с маской
- Классификатор -- нейросеть с 2 скрытыми слоями
  - Обучаем фиксированное число эпох, на каждой измеряем AUC на валидационной выборке
  - Выбираем эпоху с наилучшим AUC
- Ансамблирование
  - Настроим 30 нейросетей на  $\frac{2}{3}$  признаков, 30 на  $\frac{1}{2}$
  - Поделим выборку с признаками на обучение и контроль по-другому, повторим
  - Смешивание -- простое усреднение ~ 150 сетей



# Другая нейросеть

- Можно попробовать обучить собственную сеть по данным задачи.
- Некоторые нейроны учатся находить текст (на изображении справа показаны активации нейронов при прогоне изображения слева через нейросеть)



# Основные результаты

- Исследовано несколько подходов к решению задачи
- Показано, что признаки ImageNet сети могут быть использованы в задаче поиска контактной информации на изображениях
- Итоговое качество алгоритма на выборке C составило ~ 93.7% AUC