

Как найти и закрыть гос. номер на фото автомобиля и помешать копированию контента с помощью adversarial attack

Сергеев Илья



UseData
Conf
2019

Профессиональная конференция
для специалистов по машинному
обучению и анализу данных



@sergeevii123

Senior DS in Avito
команда: DS as a service

что я делаю – computer vision

что вообще делаем:

CV

OCR

NLP





HABR.COM

Для чего и как мы скрываем госномера автомобилей в объявлениях Авито



HABR.COM

Как мы боремся с копированием контента, или первая adversarial attack в проде



HABR.COM

Как мы боремся с копированием контента, или первая adversarial attack в проде

204 лайка (топ 1 в блоге Авито)
> 40 к просмотров



HABR.COM

Для чего и как мы скрываем госномера автомобилей в объявлениях Авито

69 лайков
> 50 к просмотров

181 комментарий
(из них ~10 технические)



HABR.COM

Для чего и как мы скрываем госномера автомобилей в объявлениях Авито



razielvamp 10 апреля 2019 в 07:38



+9



А что касается статьи, то я зашел почитать про «Для чего» — прочитав по диагонали, внятного ответа не нашел.

[Ответить](#)




Зачем?

1. По гос. номеру можно найти много дополнительной информации о машине

-  **GLeBaTi** 9 апреля 2019 в 13:51 # 📌 ↑ +36 ↓

Так и не ясно для чего скрывать номера. Мошенники как-то могут воспользоваться этой информацией?

[Ответить](#)

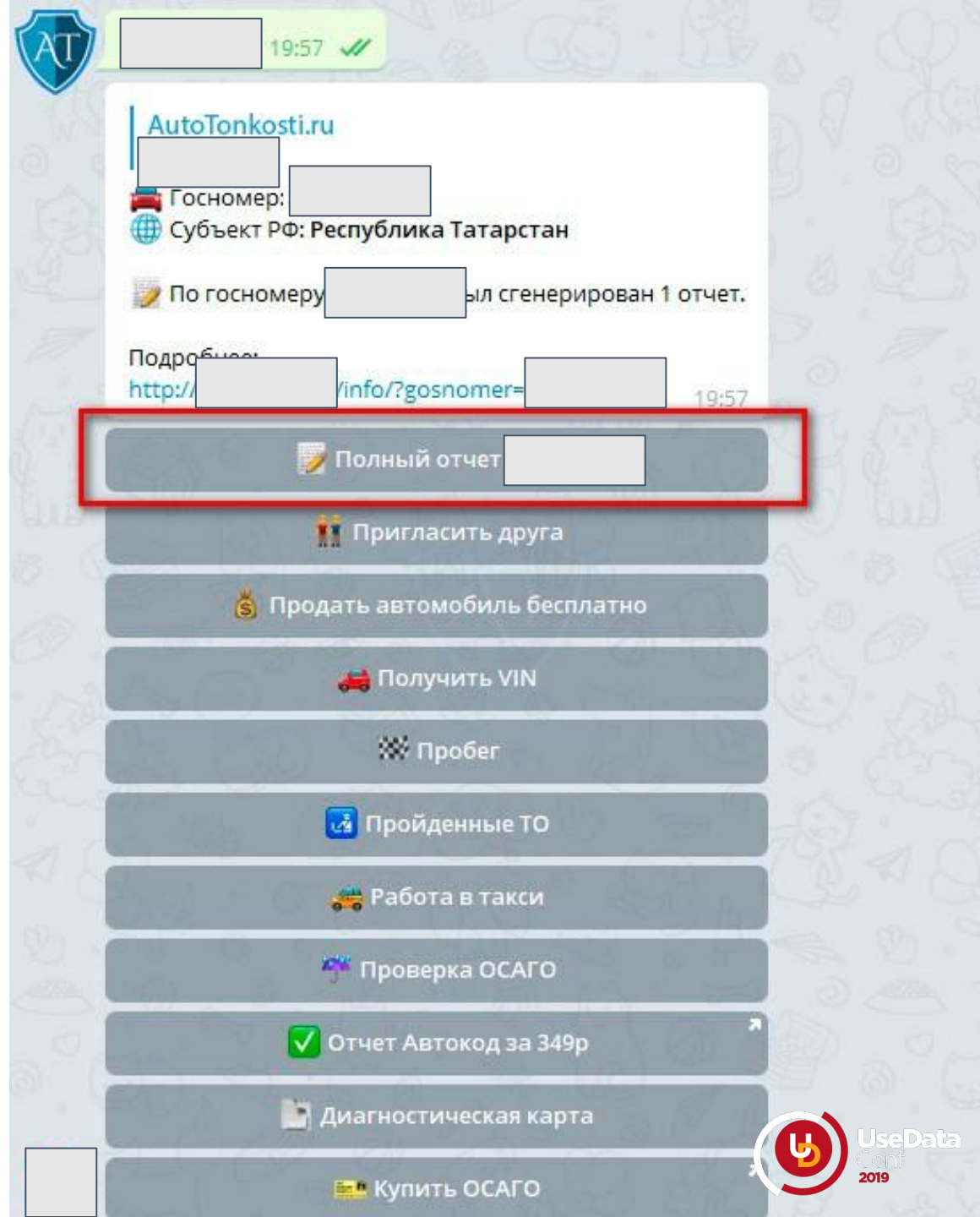
- •  **Cubist** 9 апреля 2019 в 13:58 # 📌 🗨️ 📄 ↑ -7 ↓

Могут воспользоваться. По ним можно узнать и имя собственника, контакты и место проживания, паспортные данные и тд.

bot в телеграме

Госномер: В [REDACTED] TM74
Машина: BA321083
Дата постановки на учет: 18.02.2005
Год машины: 1990
Владелец: [REDACTED] НАДЕЖДА
АНАТОЛЬЕВНА
Паспорт: 75 [REDACTED] 17, ЧЕЛЯБИНСК
БАРБЮСА д.140 [REDACTED]
Год рождения владельца: 1960
СТС: 74Н [REDACTED] 29 (АА [REDACTED] 786)
ПТС: 74Е [REDACTED] 2170

11:37



Зачем?

1. По гос. номеру можно найти много дополнительной информации о машине
2. Некоторые пользователи Авито сами закрывают гос. номер







Зачем скрывать гос. номер на фото авто?

1. По гос. номеру можно найти много дополнительной информации о машине
2. Некоторые пользователи Авито сами закрывают гос. номер
3. Наши конкуренты уже это сделали

Масштаб проблемы

В день 20 000 новых объявлений в Авто

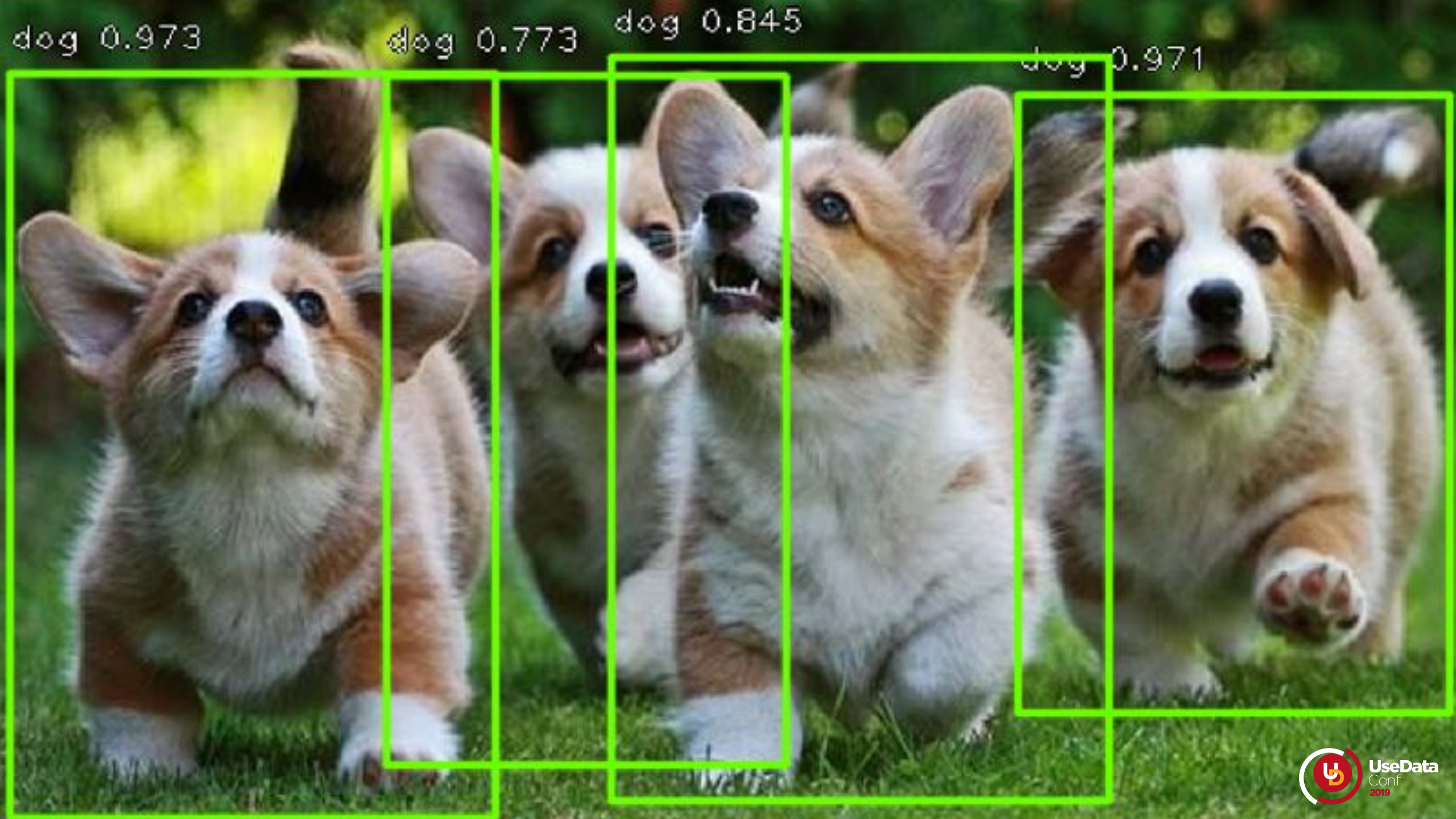
За 2018 год было продано 2,5 миллиона автомобилей ~7000 в день

dog 0.973

dog 0.773

dog 0.845

dog 0.971

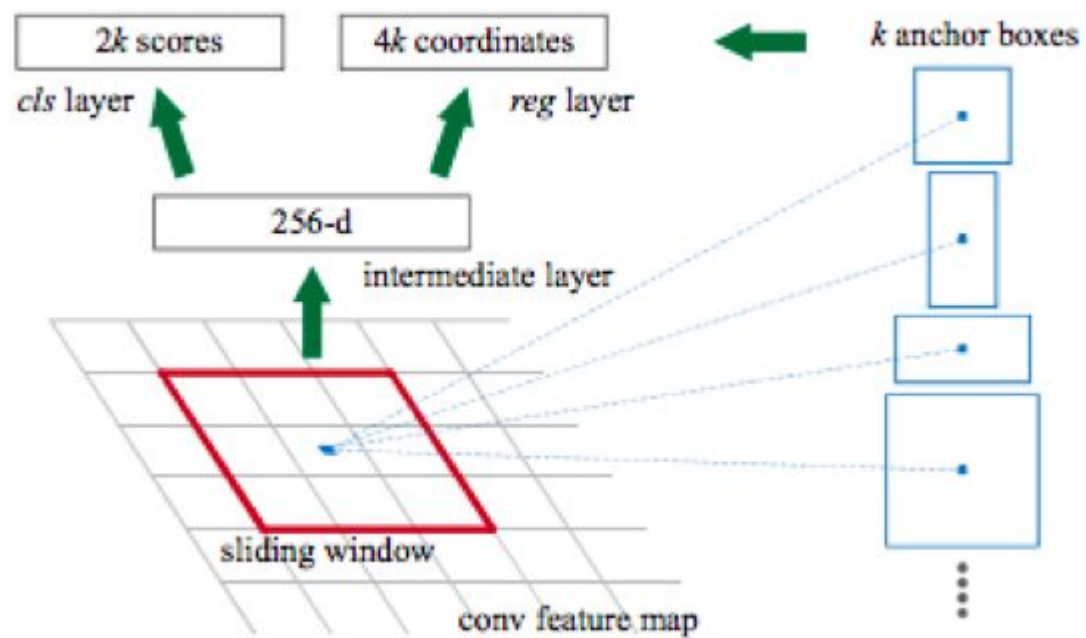


Object detection

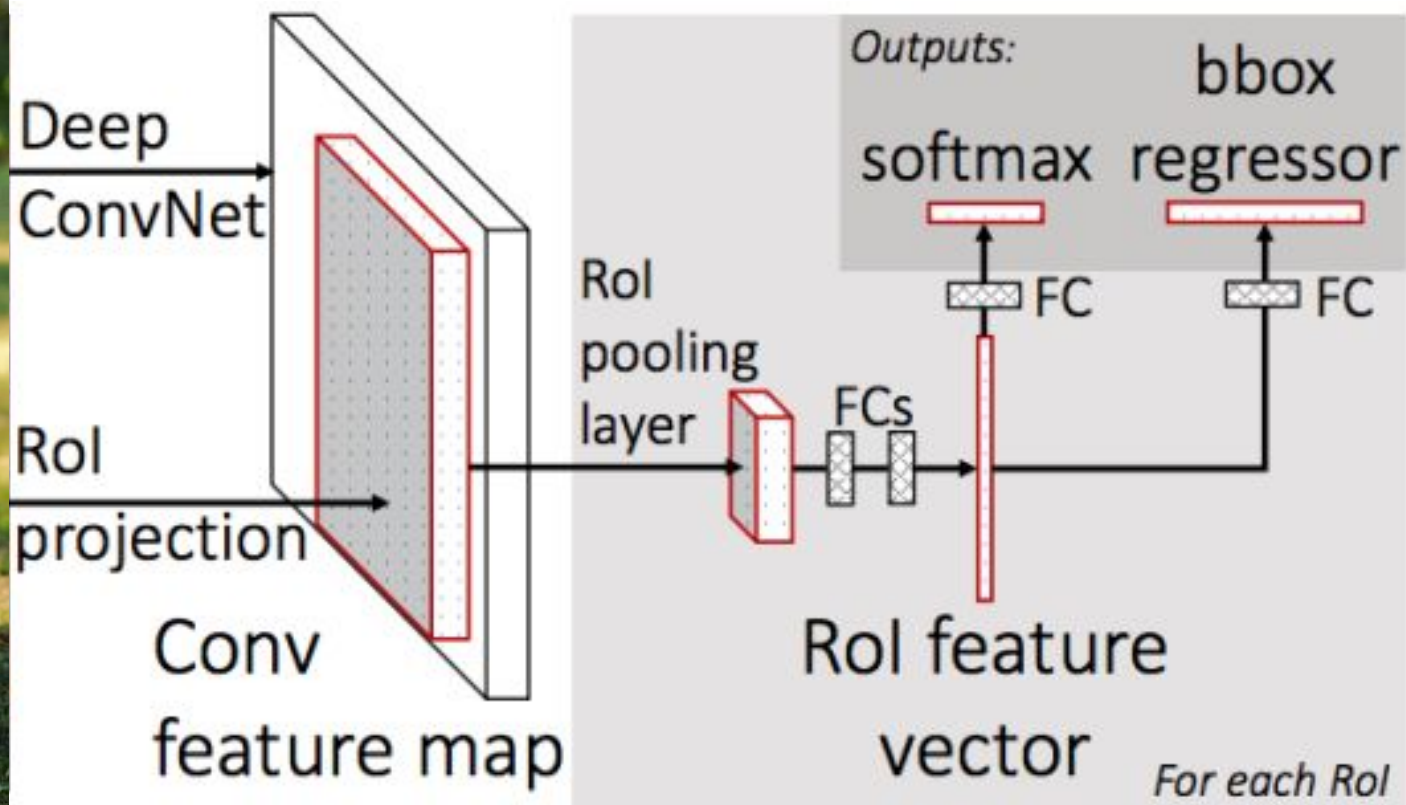
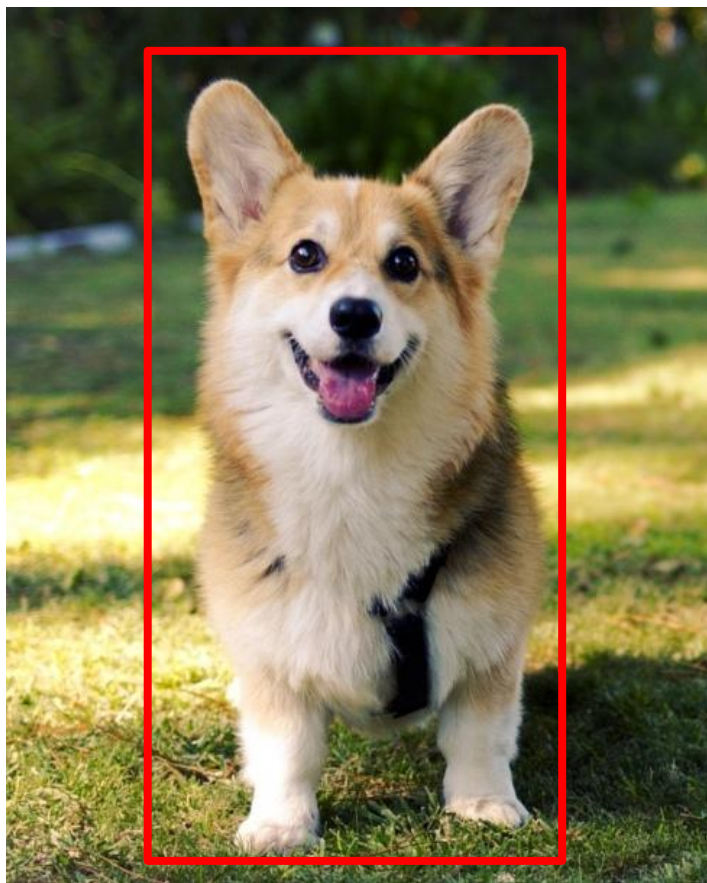
Двухэтапные модели
Faster RCNN, Mask
RCNN

Одноэтапные модели
SSD, YOLO, RetinaNet

Двухэтапные детекторы



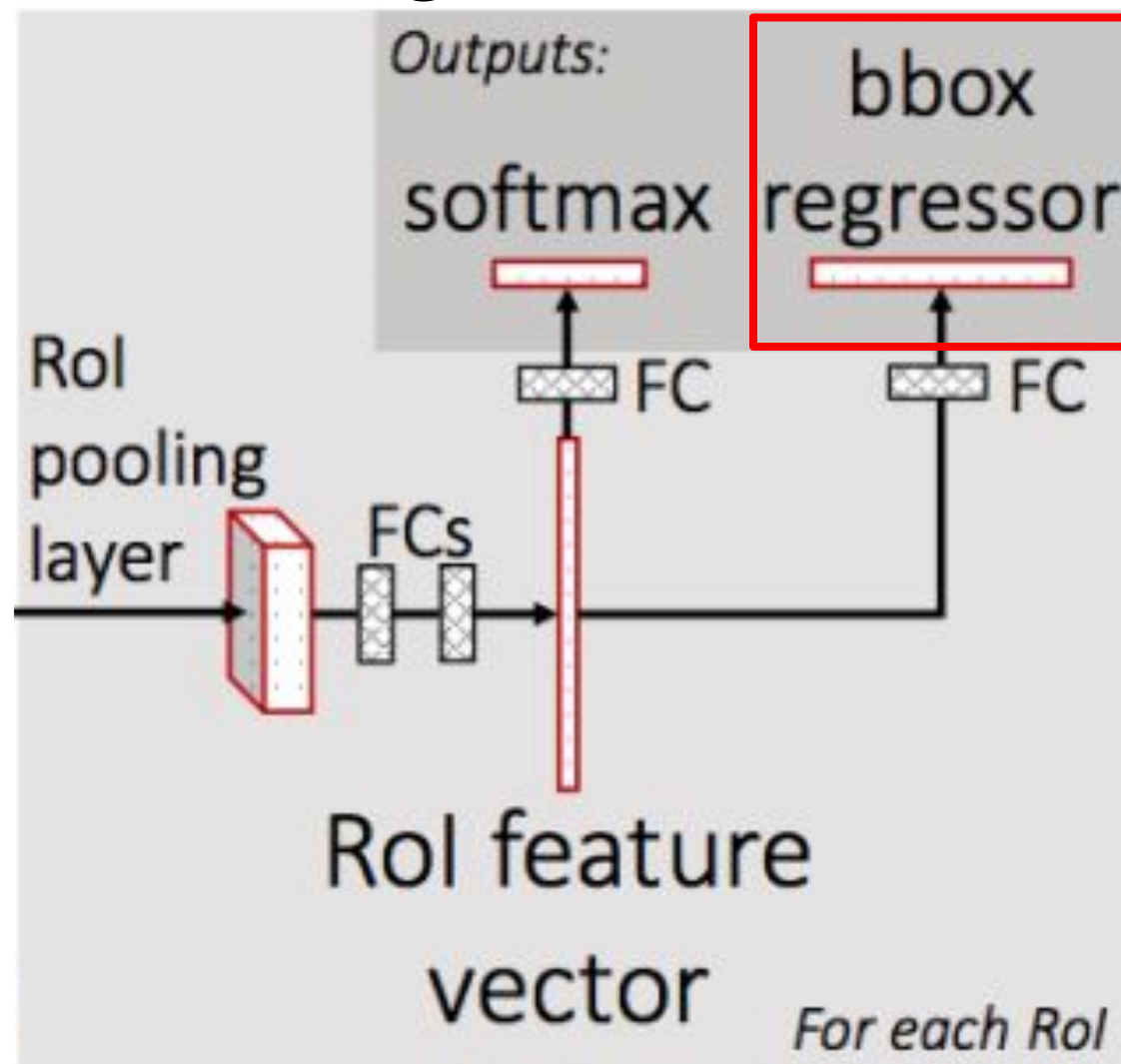
Двухэтапные детекторы





 Avito

Изменить bbox regressor



Нужна ли тяжелая сеть?

1. Бинарная классификация

Нужна ли тяжелая сеть?

1. Бинарная классификация
2. На фото один номерной знак

Нужна ли тяжелая сеть?

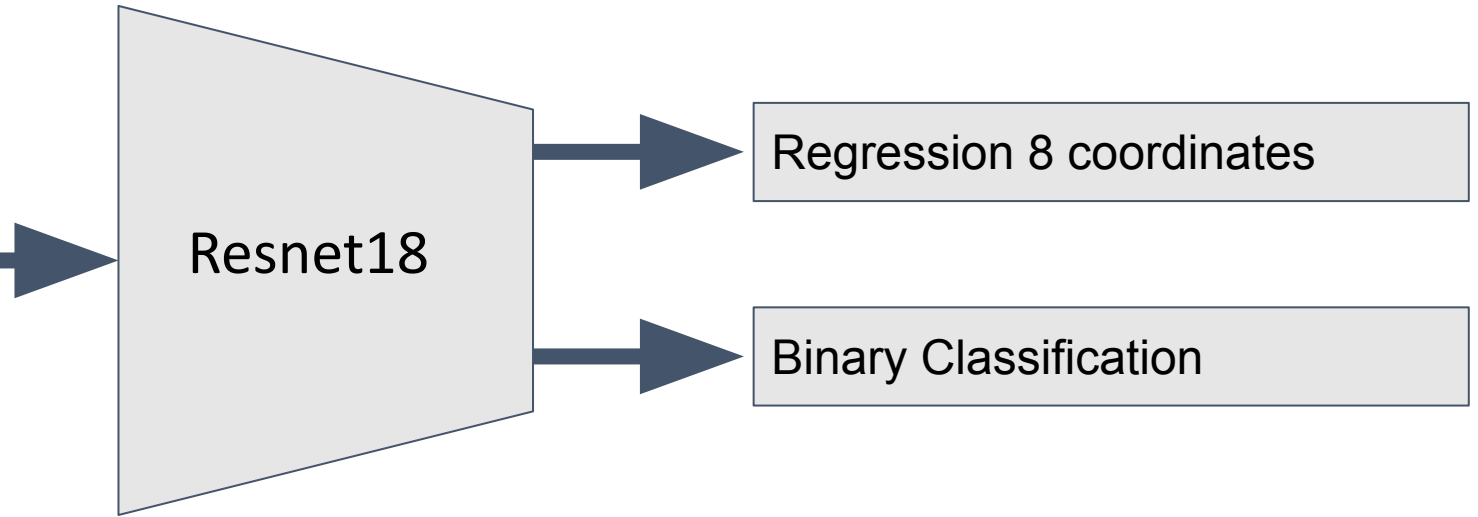
1. Бинарная классификация
2. На фото один номерной знак
3. Производительность

Нужна ли тяжелая сеть?

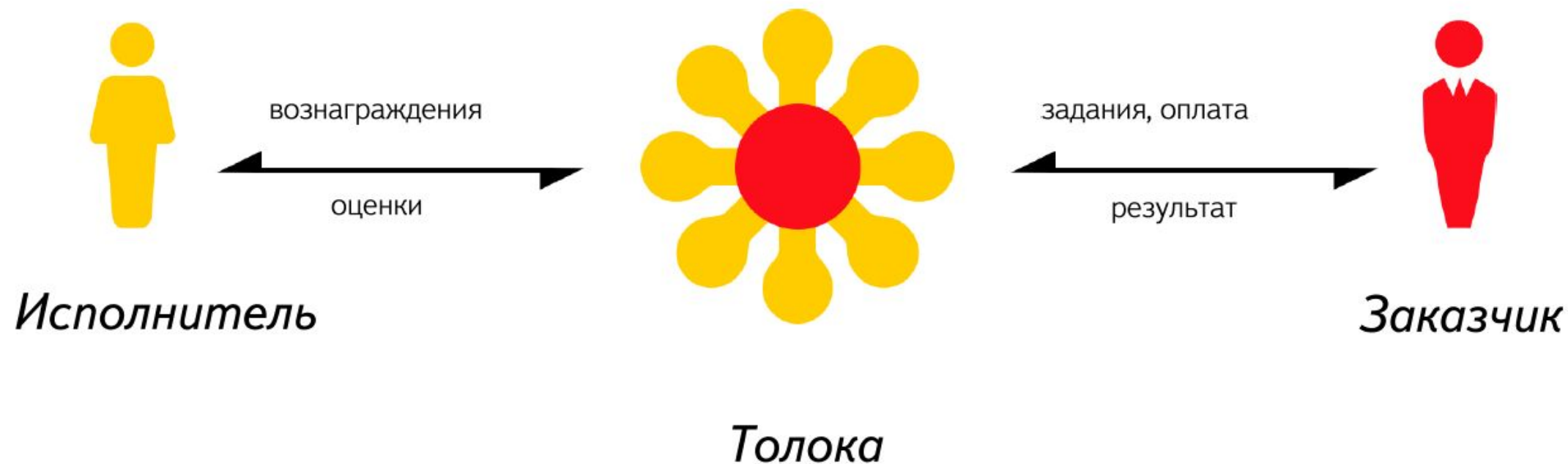
 apelsyn 10 апреля 2019 в 15:53      0 

И еще у нас медленее, все работает на Mask RCNN, в среднем, 0.8 s/фото на GPU.

[Ответить](#)



Данные





Настройка толоки

1. Правила против ботов
2. Honey-поты

СКОЛЬКО СТОИТ?

4000 картинок
перекрытие 3
всего 28\$

Обучение

$$L_{1;smooth} = \begin{cases} |x| & \text{if } |x| > \alpha; \\ \frac{1}{|\alpha|} x^2 & \text{if } |x| \leq \alpha \end{cases}$$

$$BCE = -\frac{1}{N} \sum_{i=0}^N y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)$$

Прод!

nvidia-docker
веса в git lfs
kubernetes



accuracy на тестовой выборке 0.98
95-й перцентиль 250 мс

**Toyota Corolla, E110**

150 000 ₽



Обновлено день назад

236 000 км

1.3 л, Бензин

85 л.с.

1999 г.

Хетчбэк

Механика



Алексей

Пермь, показать на карте

+7 (966) 795 [Показать номер](#)**Toyota Yaris, XP9**

260 000 ₽



Обновлено день назад

300 000 км

1 л, Бензин

69 л.с.

2008 г.

Хетчбэк

Механика



Светлана

Пермь, показать на карте

+7 (965) 555 [Показать номер](#)**Mitsubishi Pajero Pinin, 1 поколение**

175 000 ₽



Обновлено день назад

200 000 км

1.8 л, Бензин

114 л.с.

2004 г.

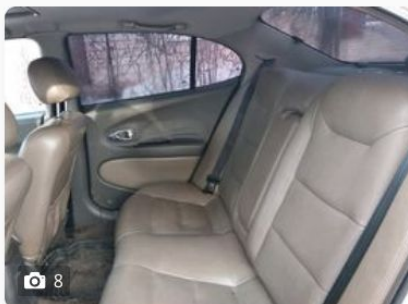
Внедорожник

Механика



Продавец

Грозный, показать на карте

+7 (989) 175 [Показать номер](#)**Daewoo Leganza, 1 поколение**

160 000 ₽



Обновлено день назад

154 000 км

2.2 л, Бензин

136 л.с.

2001 г.

Седан

Автомат



Лучшая комплектация

Наталья

Пермь, показать на карте

+7 (919) 455 [Показать номер](#)

^ Все марки, 1 параметр

49 объявлений

Сохранить



LADA (VAZ) 2109

150 000 Р

2003

75 700 км

1.5 л / 78 л.с. / Бензин Передний
Механика Чёрный
Хэтчбек 5 дв.

VIN ПРОВЕРЕН

Грозный, 1 час назад



Ford Focus III

620 000 Р

2013

74 017 км

1.6 л / 125 л.с. / Бензин Передний
Робот Белый
Универсал 5 дв.

VIN ПРОВЕРЕН ЕСТЬ ИСТОРИЯ

Грозный, 1 час назад



Ford Focus III

565 000 Р

2012

180 000 км

1.0 л / 125 л.с. / Бензин Передний
Механика Белый
Седан

VIN ПРОВЕРЕН

Грозный, 1 час назад



Toyota Camry VII (XV50)

850 000 Р

2013

175 000 км

3.5 л / 249 л.с. / Бензин Передний
Автомат Чёрный
Седан

VIN ПРОВЕРЕН

Гудермес, 1 час назад

Что происходит?



Что с картинками?



Что с картинками?





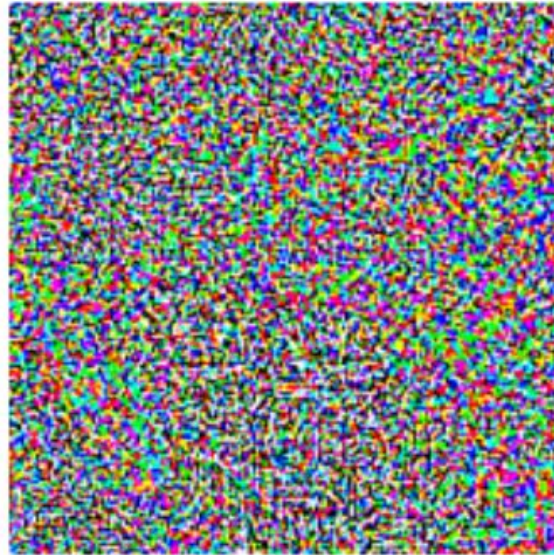
watermark

Adversarial Examples



“panda”
57.7% confidence

+ .007 ×



“nematode”
8.2% confidence

=



“gibbon”
99.3 % confidence

Examples



Examples

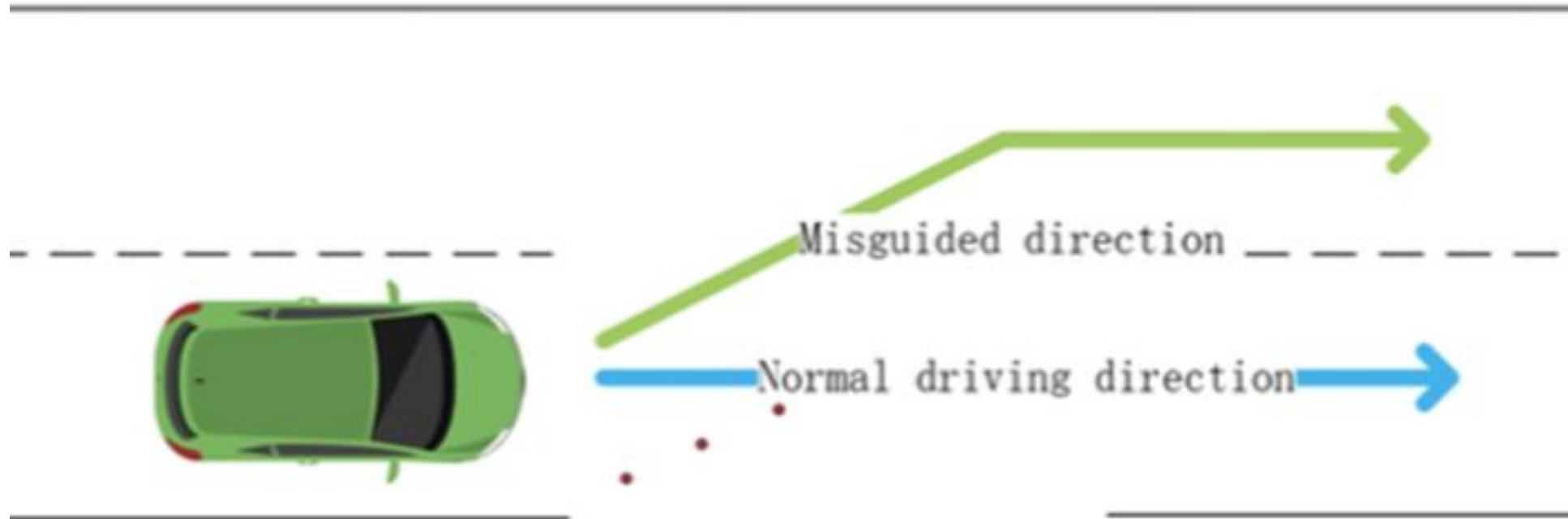
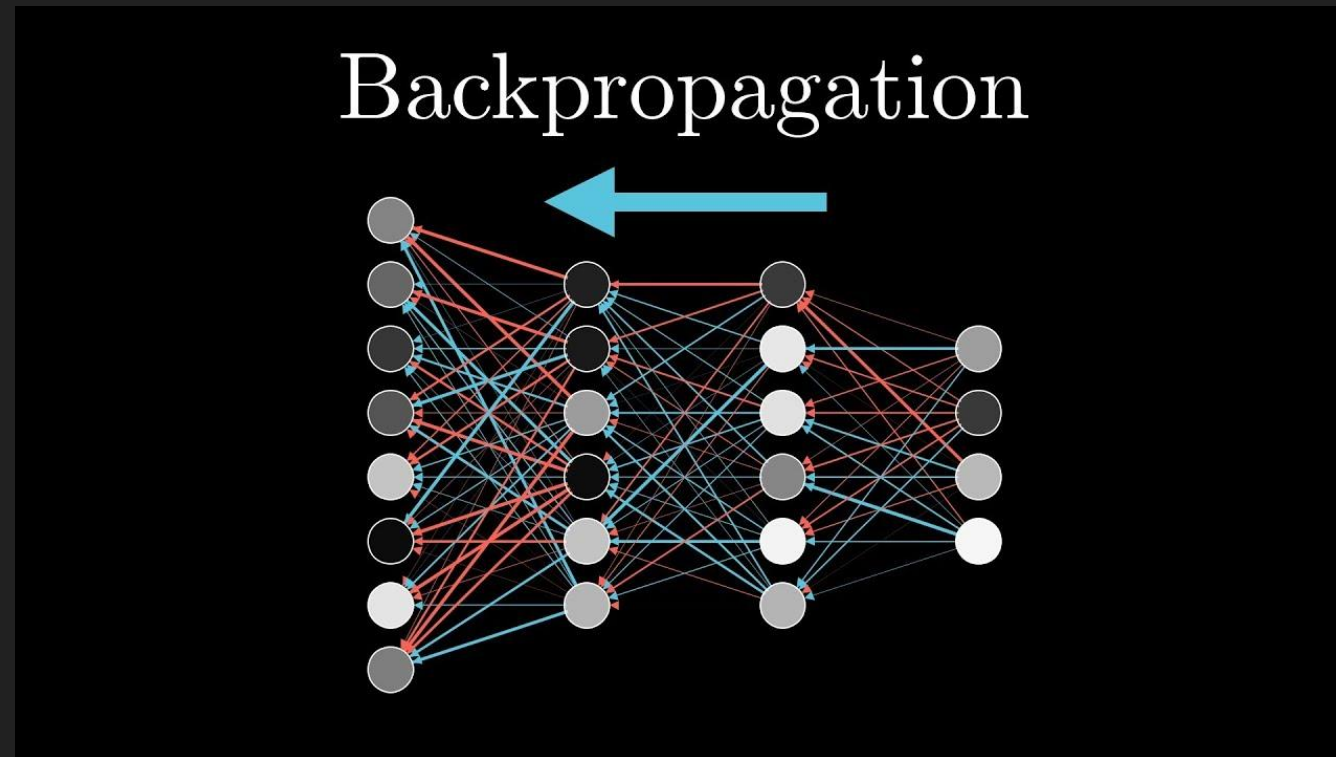


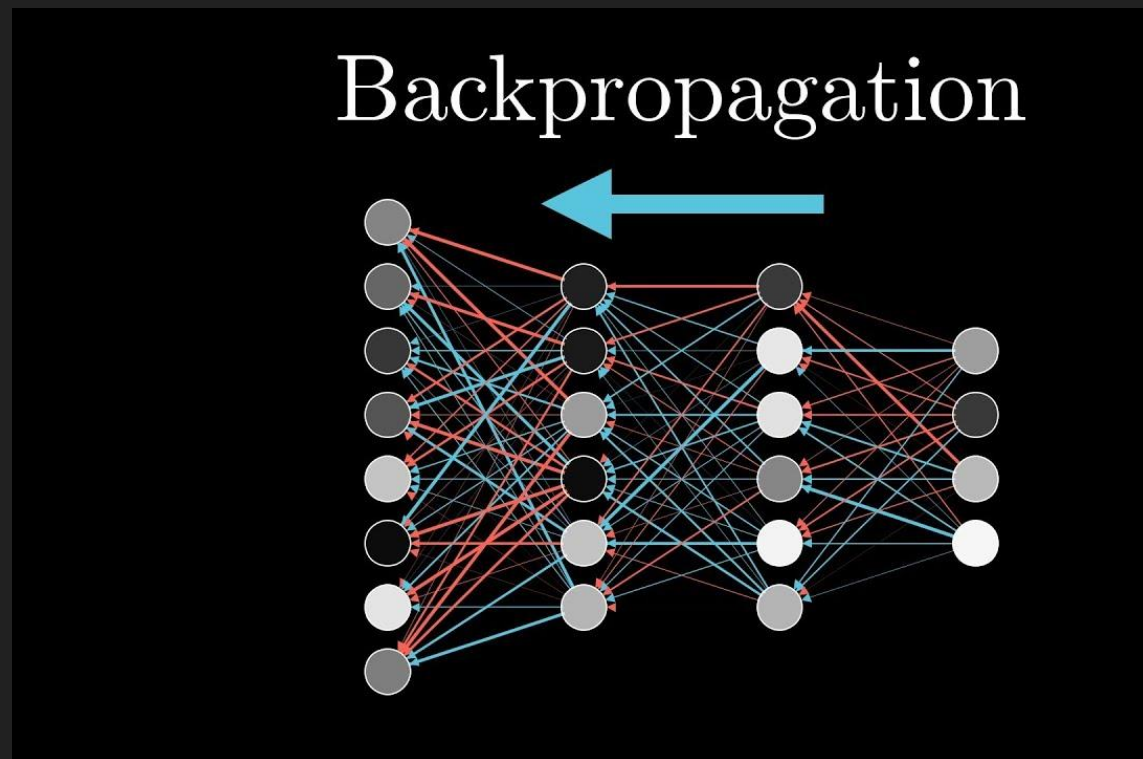
Fig 34. Fake lane mode in physical world

Visualization, Deep Dream, Neural Style, Adversarial Examples



[CS231n youtube lecture](#)

Visualization, Deep Dream, Neural Style, Adversarial Examples

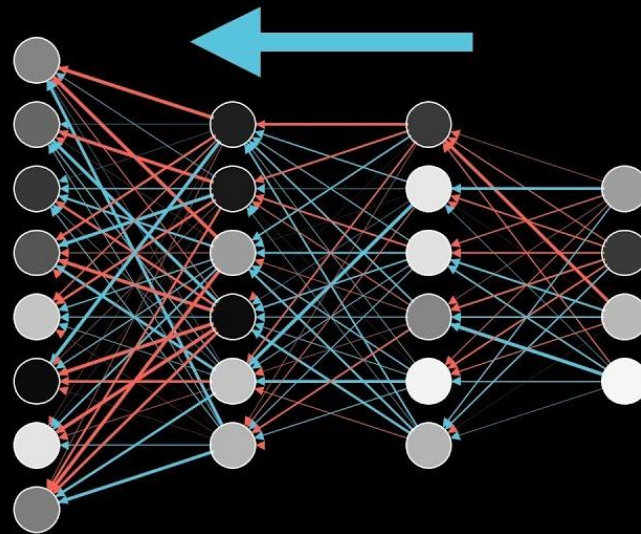


[CS231n youtube lecture](#)

Visualization, Deep Dream, Neural Style, Adversarial Examples



Backpropagation

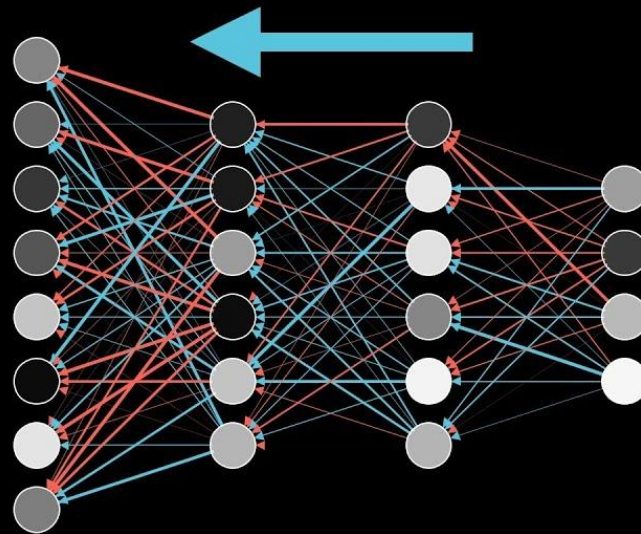


[CS231n youtube lecture](#)

Visualization, Deep Dream, Neural Style, Adversarial Examples



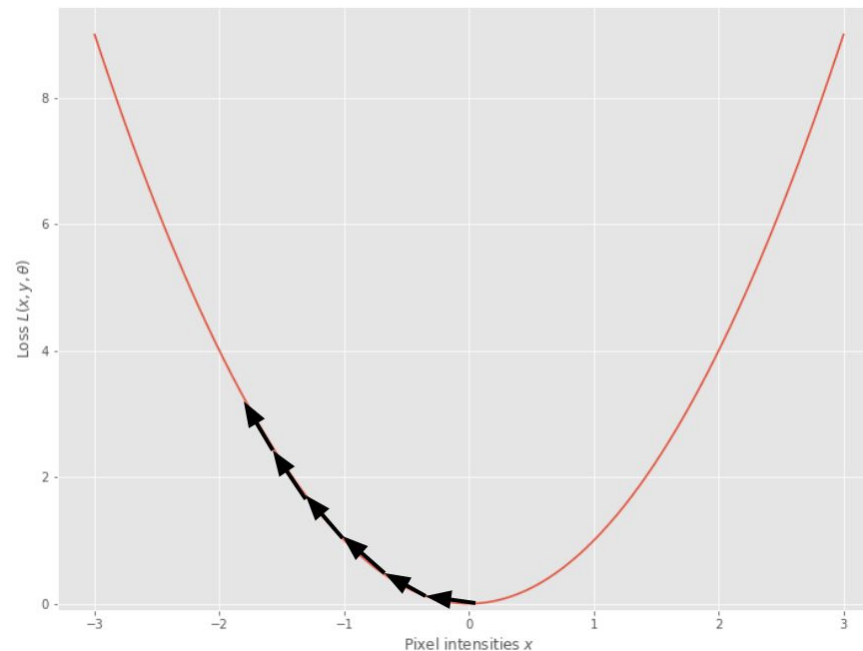
Backpropagation



[CS231n youtube lecture](#)

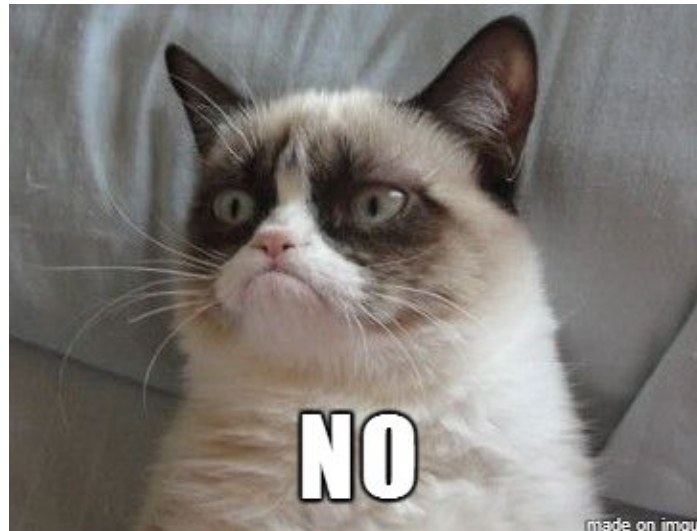
Итеративный метод

$$x_0^{adv} = x, \quad x_{t+1}^{adv} = x_t^{adv} + \alpha \cdot \text{sign}(\nabla_x J(x_t^{adv}, y)).$$



Итеративный метод

$$x_0^{adv} = x, \quad x_{t+1}^{adv} = x_t^{adv} + \alpha \cdot \text{sign}(\nabla_x J(x_t^{adv}, y)).$$



Fast gradient sign method (FGSM)

$$x^{adv} = x + \varepsilon \cdot \text{sign}(\nabla_x J(x, y_{true})),$$

where

x is the input (clean) image,

x^{adv} is the perturbed adversarial image,

J is the classification loss function,

y_{true} is true label for the input x .



Targeted fast gradient sign method (T-FGSM)

$$x^{adv} = x - \varepsilon \cdot \text{sign}(\nabla_x J(x, y_{target})),$$

where

y_{target} is the target label for the adversarial attack.



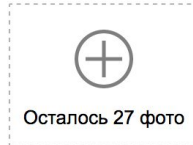
Avito 606TE 174

 **LADA (ВАЗ) / 1111 Ока / 2008 Хэтчбек 3 дв. / СеАЗ 1.0 MT (53 л.с.)**

Очистить

Фотографии и видео

Рекомендации и ограничения



Лучшие ракурсы

Автоматически расставить фотографии наиболее привлекательным образом.

Добавьте видеоролик с вашим автомобилем — это привлечёт больше внимания, повысит доверие к продавцу и увеличит вероятность звонка.

ссылка на Youtube

Добавить

Цвет автомобиля



Пробег

км

Битый или не на ходу  Не растаможен

Личные данные и место осмотра

Город продажи
Москва

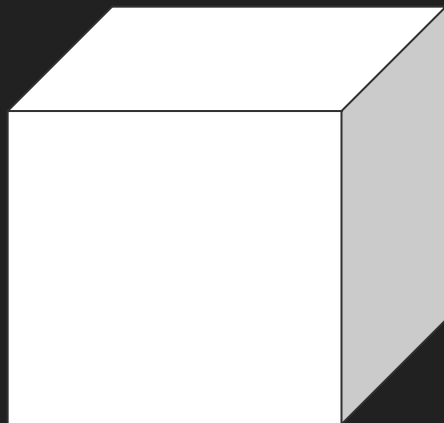
Изменить город после размещения будет нельзя

Качество объявл

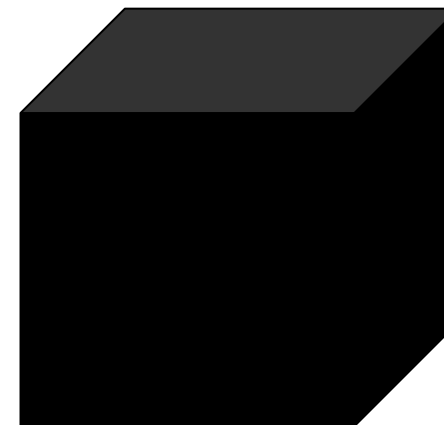
- Марка и моде
- Фото и виде
- Цвет и пробег
- Цена и контак
- Состояние
- Комплектаци
- Описание

39%

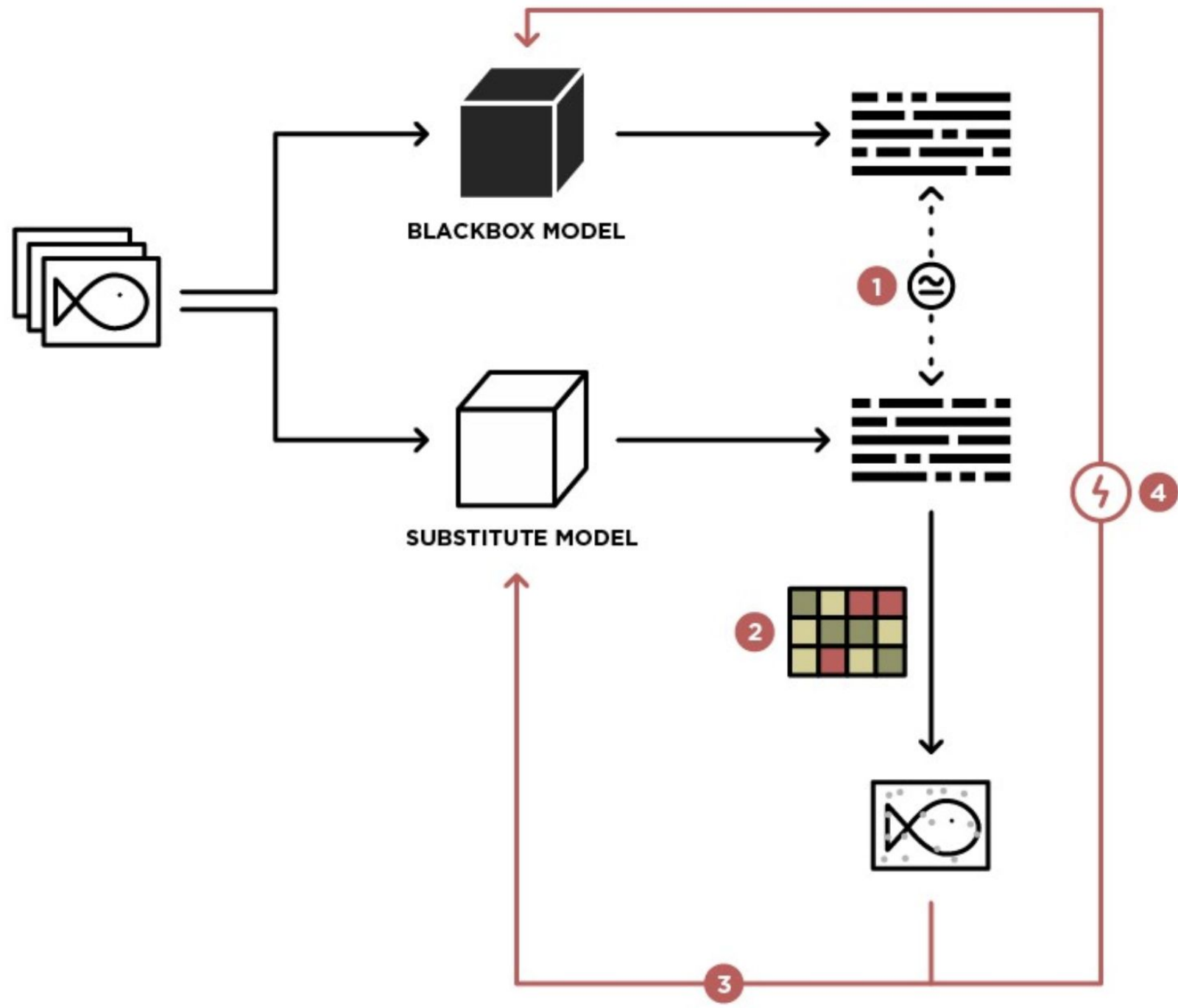
White box vs Black box



- Архитектура сети известна
- Гиперпараметры известны
- Можно получить предсказания и градиент



- Архитектура сети неизвестна
- Гиперпараметры неизвестны
- Можно получить предсказания (с ограничениями)



ShapeShifter: Robust Physical Adversarial Attack on Faster R-CNN Object Detector



(a) Person (low)



(b) Sports ball (low)



(c) Untargeted (low)



(d) Person (high)



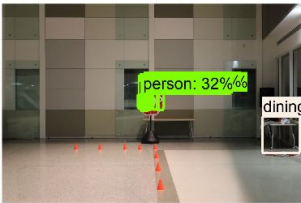
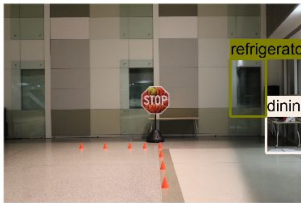

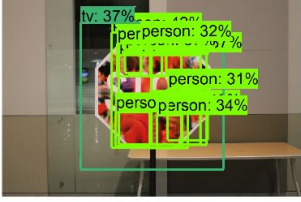



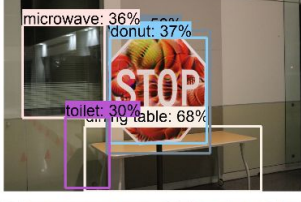
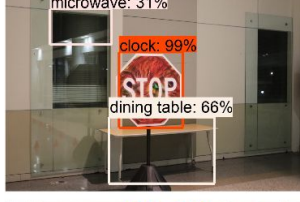
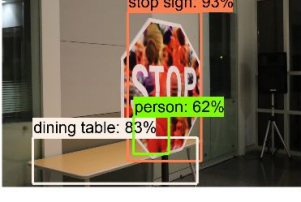
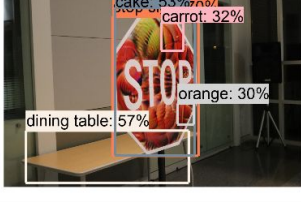
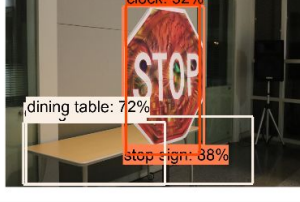
(e) Sports ball (high)



(f) Untargeted (high)

<https://arxiv.org/abs/1804.05810>

ShapeShifter: Robust Physical Adversarial Attack on Faster R-CNN Object Detector

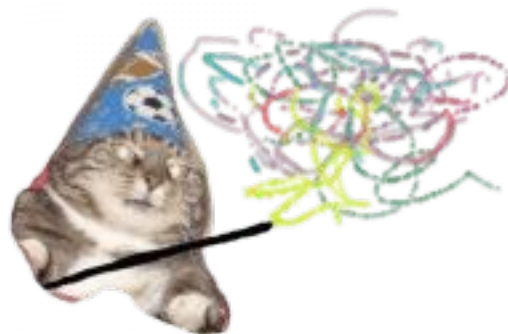
Dist.	Angle	Target: person	Target: sports ball	Untargeted
40'	0°	 person: 32% dining	 refrigerator dining	 stop sign: 30% stop
10'	0°	 tv: 37% person: 32% person: 31% person: 34%	 cake: 36% STOP	 clock: 99% STOP
10'	30°	 refrigerator: 46% person: 36% person: 31% person: 35% person: 42% dining table: 73%	 microwave: 36% donut: 37% STOP toilet: 30% table: 68%	 microwave: 31% clock: 99% dining table: 66% STOP
5'	60°	 stop sign: 93% person: 62% dining table: 83%	 cake: 53% carrot: 32% orange: 30% dining table: 57% STOP	 clock: 52% dining table: 72% stop sign: 88% STOP

<https://arxiv.org/abs/1804.05810>

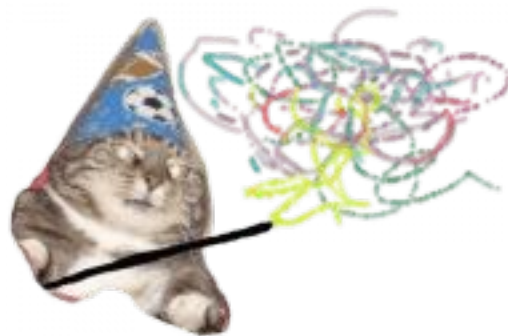
Насколько переобучена сеть автору?



Насколько переобучена сеть автору?



Насколько переобучена сеть автору?









 auto.ru



Avito





 auto.ru

Продажа Volkswagen > Jetta > V > Седан > 1.6 MT (102 л.с.) > в Артёмовском

Volkswagen Jetta V

350 000 ₺

22 января 👁 542 (10 сегодня) № 1083480258

от 7 500 ₺/мес



Volkswagen Jetta. Подбор л
100% готовность к продаже. Проверьте



Андрей
Артёмовский



Написать

Показать телефон

+7 ●●●●●●●●

Год выпуска	2008
Пробег	159 000 км
Кузов	Седан
Цвет	Чёрный
Двигатель	1.6 л / 102 л.с. / Бензин
Коробка	Механическая
Привод	Передний
Руль	Левый
Состояние	Не требует ремонта
Владельцы	3 или более
ПТС	Оригинал
Таможня	Растаможен
VIN	XW8ZZZ1K*8G****18

[Характеристики модели в каталоге](#)



Кредит на это авто!
Первый взнос 0%. Без КАСКО!

Банк-партнер: ПАО «Совкомбанк» Лицензия Е



Продажа Opel > Corsa > C Рестайлинг > Хэтчбек 5 дв. > 1.2 АМТ (80 л.с.) > в Самаре

Opel Corsa C Рестайлинг

200 000 ₹ ▾

22 января 👁 244 (5 сегодня) № 1083479076

от 4 300 ₹/мес



Спецпредложения на OPEL
 Распродажа авто 2018 г.в! До 31.01.19 Реклама



максим
Самара



Написать

Показать телефон

+7 ●●●●●●●●

Год выпуска	2005
Пробег	180 000 км
Кузов	Хэтчбек 5 дв.
Цвет	Серебристый
Двигатель	1.2 л / 80 л.с. / Бензин
Коробка	Роботизированная
Привод	Передний
Руль	Левый
Состояние	Не требует ремонта
Владельцы	2 владельца
ПТС	Оригинал
Таможня	Растаможен
VIN	W0L0XCF6*66****97



[Характеристики модели в каталоге](#)



Kia Spectra 1.6 MT
(101 л.с.)

225 000 руб



Opel Corsa 1.2 AT
(80 л.с.)

200 000 руб



Nissan X-Trail 2.0
MT (140 л.с.) 4WD

350 000 руб



АВТО.ру

1st adversarial attack in prod

YOU'RE ABOUT
TO HACK TIME,
ARE YOU SURE?

YES NO

Imagenet-trained CNNs are biased towards texture

