



Московский государственный университет имени М.В. Ломоносова

Факультет вычислительной математики и кибернетики

Кафедра интеллектуальных информационных технологий

Лаборатория компьютерной графики и мультимедиа

Гончаренко Дмитрий Александрович

Алгоритм изменения времени суток на изображении

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Научный руководитель:

К.С. Зипа

Москва, 2019

Аннотация

Алгоритм изменения времени суток на изображении относится к классу задач машинного обучения по *переносу изображений*¹. Данная сфера значительно продвинулась благодаря современным вычислительным возможностям, в частности переносу обучения на графические процессоры, GPU. За последние несколько лет появилось немало исследовательских работ на тему переноса изображений, стилей и колоризации. В данной работе рассматриваются современные подходы к переносу изображений на примере изменения времени суток на изображении. Проводится описание нейросетевых моделей и сравнительный анализ серии экспериментов обучения.

Abstract

The algorithm of changing the time of the day on images is a subclass of Machine Learning problems of image translation. This area has advanced significantly due to the modern computing capabilities, in particular the training transfer on GPUs. Over the past few years, many research papers have appeared on the subjects of images translation, styles transferring and colorization. This research reveals modern approaches of image translation on the example of changing the time of day on the image. A description of the neural network models and a comparative quality analysis of a series of training experiments are carried out.

¹ **Перенос изображений** (англ. *image translation*, или *image transferring*) – подвид технологии переноса обучения, позволяющий сохранять и объединять локальные признаки изображений.

Содержание

1	Введение	3
2	Предметная область	4
2.1	Обучение с учителем	4
2.2	Обучение без учителя	4
3	Постановка задачи	5
3.1	Цель работы	5
3.2	Решаемые задачи	5
3.3	Формальная постановка задачи	5
3.4	Обзор существующих методов	6
4	Список литературы	7

1 Введение

Математические описания моделей машинного обучения появились еще в середине 20-го века. А первые попытки их практической реализации начались в конце 50х годов. В наше время задачи машинного обучения все также не теряют своей актуальности. Рост числа работ, открытие исследовательских центров внутри компаний и при университетах показывает высокий интерес к данной сфере не только в науке. Появление удобных и современных инструментов для построения самых разнообразных моделей обучения, возможности мгновенно обмениваться информацией – оказало немалое влияние на развитие *науки о данных*¹.

Стоит отметить, что не последнюю роль здесь сыграли многократно выросшие вычислительные возможности. Это позволило обучать *нейронные сети*² на персональных компьютерах за доли секунды, для чего ранее могли требоваться дни, а то и недели на специализированных вычислительных устройствах.

Перенос обучения является одной из центральных исследовательских задач современного машинного обучения. Данная область направлена на получение некоторой информации об объекте, сохранение и последующее применение этих знаний к другому объекту, связанному с первым. *Перенос изображений*, в свою очередь, является подклассом переноса обучения применяемым на изображениях. В настоящее время перенос изображений используется самыми разнообразными способами. С помощью него можно добиться объединения стилей двух изображений [1], *колоризации*³ черно-белых фотографий [2], объединение локальных признаков объектов и животных [3], увеличение разрешения изображений [4].

В данной исследовательской работе рассматриваются и систематизируются современные подходы переноса изображений, на примере изменения времени суток на изображениях. А также проводится обучение моделей на различных наборах данных и сравнительный анализ полученных результатов с использованием предложенной метрики.

В качестве инструмента был выбран фреймворк PyTorch⁴, так как он разработан в виде библиотеки для языка Python - основного языка современного машинного обучения [10], и зарекомендовал себя мощным и гибким инструментом для исследования и обучения нейронных сетей.

¹ **Наука о данных** (англ. *data science*) – соовокупность процессов и методов направленных на извлечение информации из исследуемых данных.

² **Нейронная сеть** – математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей – сетей нервных клеток живого организма. Является мощным современным инструментом машинного обучения. Главная особенность – способность обучаться на предоставленных данных, называемых тренировочными. Помимо исследовательских направлений нейронные сети также активно используются в коммерции, например обработка спама на электронной почте и система рекомендаций в интернет-магазинах.

³ **Колоризация** (англ. *colorization*) – преобразование монохромных изображений в цветные.

⁴ **PyTorch** – библиотека машинного обучения, основанная на Torch. Разработана исследовательской группой в Facebook.

2 Предметная область

2.1 Обучение с учителем

Обучение с учителем (англ. *supervised learning*) – способ машинного обучения, в ходе которого входные объекты соотносятся с выходными до начала обучения. Обучение происходит на заготовленных парах объектов. Задача метода – построить функцию внутренней зависимости между примерами, которая затем будет отображать входные изображения желаемым образом.

2.2 Обучение без учителя

3 Постановка задачи

3.1 Цель работы

В науке о данных одну из ключевых ролей играет непосредственный выбор данных, на которых будет обучаться сеть, и которые будут использоваться для валидации качества модели. Перенос изображений не является исключением. Цель работы – получить модель, которая будет уметь с высокой точностью переносить изображения между *доменами*¹ и способную решить проблему изменения времени суток на изображениях с приемлемой точностью.

3.2 Решаемые задачи

Проблему изменения времени суток на изображении можно разбить на следующие этапы:

- Исследование существующих методов переноса изображений и анализ решений родственных задач
- Поиск применимых к проблеме тренировочных данных – изображений с различными временами суток, и их применение
- На основе проведенного исследования и обзора литературы, выбор модели, подходящей для решения поставленной задачи наилучшим образом
- Множественное обучение выбранной модели на различных конфигурациях с помощью библиотеки PyTorch
- Разработка и применение метрики качества обучения для выявления наилучших параметров сети
- Проведение серии экспериментов переноса изображений для получения оценки качества и скорости работы модели

3.3 Формальная постановка задачи

Формально алгоритм изменения времени суток на изображении можно сформулировать следующим образом:

Вход

На вход алгоритму поступают два изображения из различных доменов $x_1 \in X_1$ и $x_2 \in X_2$, где

- X_1 – домен с набором изображений первого типа
- X_2 – домен с набором изображений второго типа

Не ограничивая общности, пусть X_1 содержит изображения дня, а X_2 соответствует ночи.

Выход

Необходимо построить *кроссдоменные*² отображения $f_{12}: X_1 \rightarrow X_2$ и $f_{21}: X_2 \rightarrow X_1$.

Реализовать функцию, принимающую на вход изображения из разных доменов x_1, x_2 и возвращающую перенесенные изображения x'_2, x'_1 соответственно:

$$x'_2, x'_1 = \text{CrossDomainTranslator}(x_1, x_2), \text{ где } x'_1 \in X_1, x'_2 \in X_2 \quad (1)$$

¹ Домен – область, множество, содержащее в себе объекты одного типа.

² Кроссдоменный перенос (англ. *cross-domain transfer*) – перенос объектов между доменами

3.4 Обзор существующих методов

4 Список литературы

- [1] **J.Johnson, A.Alahi, L.Fei-Fei.** *Perceptual Losses for Real-Time Style Transfer and Super-Resolution.* Department of Computer Science, Stanford University, March 2016.
- [2] **R.Zhang, J.-Y.Zhu, P.Isola, X.Geng, A.S.Lin, T.Yu, A.A.Efros.** *Real-Time User-Guided Image Colorization with Learned Deep Priors.* University of California, Berkeley, May 2017.
- [3] **J.-Y.Zhu, T.Park, P.Isola, A.A.Efros.** *Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.* Berkeley AI Research (BAIR) laboratory, UC Berkeley, November 2018.
- [4] **Y.Yuan, S.Liu, J.Zhang, Y.Zhang, C.Dong, L.Lin** *Unsupervised Image Super-Resolution using Cycle-in-Cycle Generative Adversarial Networks* Guangdong Key Laboratory of Intelligent Information Processing, Shenzhen University, Graduate School at Shenzhen, Department of Automation, Tsinghua University, September 2018.
- [5] **M.-Y.Liu, T.Breuel, J.Kautz.** *Unsupervised Image-to-Image Translation Networks.* NVIDIA Corporation, July 2018.
- [6] **X.Huang, M.-Y.Liu, S.Belongie, J.Kautz.** *Multimodal Unsupervised Image-to-Image Translation.* Cornell University, NVIDIA Corporation, August 2018.
- [7] **J.-Y.Zhu, R.Zhang, D.Pathak, A.A.Efros.** *Toward Multimodal Image-to-Image Translation.* Berkeley AI Research, UC Berkeley, Adobe Research, October 2018.
- [8] **L.Ma, X.Jia, S.Georgoulis, T.Tuytelaars, L.V.Gool.** *Exemplar Guided Unsupervised Image-to-Image Translation.* Berkeley AI Research, UC Berkeley, Adobe Research, March 2019.
- [9] **С.Николенко, А.Кадурын, Е.Архангельская.** *Глубокое обучение.* СПб.: Питер, 2018.
- [10] **The State of the Octoverse: machine learning** – GitHub: Report, January 2019.