

Автоматическая колоризация изображений на основе свёрточных нейронных сетей

Л. В. Серебряная, к. т. н., доцент, доцент кафедры программного обеспечения информационных технологий

E-mail: L_silver@mail.ru

ORCID ID: 0000-0001-7189-7378

Белорусский государственный университет информатики и радиоэлектроники, ул. П. Бровки, д. 6, 220013, г. Минск, Республика Беларусь

В. В. Потараев, м. т. н., аспирант кафедры программного обеспечения информационных технологий

E-mail: vic229@rambler.ru

ORCID ID: 0000-0002-1531-047X

Белорусский государственный университет информатики и радиоэлектроники, ул. П. Бровки, д. 6, 220013, г. Минск, Республика Беларусь

Аннотация. Выполнен анализ методов и средств колоризации изображений. Обоснован выбор модели искусственной нейронной сети для обработки графической информации. Сформулирована задача автоматической колоризации произвольных изображений. Приведены исходные данные, условия и ограничения, необходимые для модели колоризации. В результате классификации изображений получен набор гиперколонок нейронной сети для каждого обрабатываемого изображения. Создана модель колоризации, которая позволяет по набору гиперколонок определить цвет каждого пикселя изображения. Данная модель состоит из двух связанных между собой частей: классификатора и колоризатора. Классификатор основан на использовании свёрточной нейронной сети, а колоризатор – на использовании хеш-таблицы, хранящей соответствие гиперколонок и цветов. Предложен алгоритм применения данной модели для колоризации изображений. Выполнено сравнение результата колоризации для разработанной и существующей моделей. Создано программное средство, которое позволяет осуществлять обучение различных нейронных сетей и колоризацию графической информации. Эксперименты показали, что разработанная модель достаточно корректно определяет цвет изображения. Предложенный алгоритм позволяет использовать свёрточную нейронную сеть для раскрашивания черно-белых изображений, цветокоррекции цветных рисунков и др.

Ключевые слова: искусственная нейронная сеть, свёртка, классификация данных, колоризация изображений, гиперколонки

Для цитирования: Серебряная, Л. В. Автоматическая колоризация изображений на основе свёрточных нейронных сетей/ Л. В. Серебряная, В. В. Потараев// Цифровая трансформация. – 2020. – № 2 (11). – С. 58–64. <https://doi.org/10.38086/2522-9613-2020-1-58-64>



© Цифровая трансформация, 2020

Automatic Image Colorization Based on Convolutional Neural Networks

L. V. Serebryanaya, Candidate of Sciences (Technology), Associate Professor, associate professor of Information Technologies Software Department

E-mail: L_silver@mail.ru

ORCID ID: 0000-0001-7189-7378

Belarusian State University of Informatics and Radioelectronics, 6 P. Brovka st., 220013, Minsk, Republic of Belarus

V. V. Potaraev, Master of Science (Technology), Ph.D. Student of Information Technologies Software Department

E-mail: vic229@rambler.ru

ORCID ID: 0000-0002-1531-047X

Belarusian State University of Informatics and Radioelectronics, 6 P. Brovka st., 220013, Minsk, Republic of Belarus

Abstract. Analysis of methods and tools for image colorization was performed. It was explained why artificial neural network model was chosen for graphics information processing. The task of automatic colorization of arbitrary images was formulated. Initial data, conditions and constraints necessary for colorization model are listed. As a result of text classification, set of neural network hypercolumns was retrieved for each image processed. Colorization model was created which allows to determine color of each pixel based on hypercolumns set. In fact, this model consists of two related parts: classifier and colorizer. Classifier is based on using convolutional neural network, and colorizer is based on hash table which stores mapping of hypercolumns and colors. Algorithm of using this model for image colorization is proposed. Comparison of colorization results for developed and existing models was performed. Software tool was created which allows to perform learning of different neural networks and colorization of graphical information. Experiments shown that developed model determines image color quite correctly. Proposed algorithm allows to use convolutional neural network for colorizing black-and-white images, for color correction of pictures, etc.

Key words: artificial neural network, convolution, data classification, image colorization, hypercolumns

For citation: Serebryanaya L. V., Potaraev, V. V. Automatic Image Colorization Based on Convolutional Neural Networks. *Cifrovaja transformacija* [Digital transformation], 2020, 2 (11), pp. 58–64 (in Russian). <https://doi.org/10.38086/2522-9613-2020-2-58-64>

© Digital Transformation, 2020

Введение. Одно из направлений информационных технологий, связанных с обработкой изображений, относится к его колоризации. Раскрашивание черно-белых фильмов, архивных фото, эскизов, восстановление цвета после преднамеренного превращения изображения в полутоновое, цветокоррекция цветных изображений – это лишь небольшая часть задач, где колоризация нашла свое применение [1]. В связи с актуальностью задачи колоризации и ее автоматизации, в последнее десятилетие ведется активная работа по поиску новых и совершенствованию существующих методов ее решения.

Главным назначением колоризации является улучшение каких-либо характеристик изображения для его восприятия человеком или компьютером, а также извлечение полезной информации. Несмотря на использование компьютера, колоризация остается сложным и трудоемким процессом, поэтому его автоматизация является актуальной задачей.

Приведем несколько популярных подходов, применяемых для колоризации изображений: графические редакторы, цветовая маркировка, эталонное изображение. Наряду с преимуществами каждого из подходов, их общим недостатком является отсутствие автоматизации процесса выполнения. Если обратиться к обзору программных продуктов данного направления, то можно перечислить следующие: *ColorizePhoto*, *CODIY Pro*, *AKVIS Coloriage*, *Colorful Image Colorization*. Только последний автоматизирует процесс раскрашивания изображений, но не всегда обеспечивает необходимое качество получаемого результата.

В последние годы появляется всё больше работ, в которых описывается использование сверточных нейронных сетей для колоризации изображений.

Особенности автоматической колоризации изображений. Данная работа посвящена автоматической колоризации изображений. В результате анализа методов и средств колоризации в качестве математической модели была выбрана искусственная нейронная сеть (ИНС). При использовании ИНС для раскрашивания объектов на изображении не ставится задача точного определения цветов, но требуется их правдоподобие для получения реалистичной цветной копии.

Главными преимуществами модели на основе ИНС являются полное отсутствие участия пользователя и одноразовое обучение сети с последующим многократным ее использованием. К недостаткам можно отнести необходимость большого набора обучающих данных и, как следствие, медленный процесс обучения ИНС. Кроме того, невозможно обучить сеть для раскрашивания всего многообразия объектов окружающего мира.

В данном способе колоризации преобразование изображения из монохромного в цветное формируется автоматически на основе заранее отобранных примеров – это процесс обучения, который является обязательным. Применение ИНС в контексте данной задачи удобно тем, что довольно легко получить множество чёрно-белых изображений путём обесцвечивания цветных. Эти изображения можно использовать в качестве обучающей выборки. Результат на выходе ИНС сравнивается с исходным цветным изображением, после чего или продолжается обучение, или осуществляется переход к колоризации тестовых наборов изображений. Решение о готовности модели к использованию принимается при выполнении хотя бы одного из следующих условий:

– ошибка в работе сети достигла заранее определенного малого значения;

- ошибка в работе сети перестала уменьшаться;
- прошло определенное число итераций обучения.

Замечательной возможностью ИНС является ее способность к абстрагированию. Поэтому, при недостатке информации, сеть может прибегнуть к обобщению, но для этого она должна определить, какие именно объекты присутствуют на поданном на вход изображении. Для этого в данной работе предварительно решается задача отнесения изображений к определенным классам. Они классифицируются по признакам объектов, которые на них присутствуют. Признаки должны совпадать у объектов одного класса. По завершении процедуры классификации, когда появляется информация о классах объектов и множестве допустимых цветов для них, переходят к колоризации объектов на изображении.

В результате изложенного подхода можно выделить три этапа решения задачи колоризации изображений: обучение модели, её тестирование, раскрашивание тестовых образов.

Входными данными для модели колоризации являются файлы изображений в определенных форматах. На этапах обучения и тестирования модели данные представлены черно-белыми (полутоновыми) изображениями с одним каналом яркости и их цветными копиями с тремя каналами яркости (RGB). Анализ обучающей выборки позволяет определить количество классов для предложенных образов. На этапе колоризации на вход модели подаются только полутоновые изображения.

Выходными данными на этапе тестирования модели является количество успешных и ошибочных результатов в процентном соотношении, а также заключение о том, может ли предложенная модель успешно использоваться для тестовых изображений.

Выходными данными этапа колоризации являются цветные изображения с тремя каналами яркости RGB, представленные в заданных форматах.

Для обработки изображений наиболее широкое применение получили *свёрточные нейронные сети* (СНС). Одним из преимуществ данного типа сети для решения задачи колоризации является автоматическое выделение признаков разных порядков на изображении, что считается важным для успешной классификации и последующего раскрашивания изображения. Свёрточная ИНС состоит из следующего набора слоев: вход-

ной, свёрточные, нелинейные, субдискретизации, полносвязные, выходной [2].

Свёрточные слои являются основой данного типа ИНС. Они применяют необходимые фильтры к входным для них данным, выполняя операцию свёртки данных с выбранным ядром. Если свёрточных слоев несколько, каждый из них может иметь своё ядро. Сегодня в традиционной СНС принято перемежать свёрточные слои нелинейными, которые отвечают за отсечение ненужных деталей при выделении признаков.

Модель классификации изображений. Поскольку процессу колоризации изображений предшествует их классификация, выбирается подходящая для этого модель классификации образов. Она имеет следующий общий вид:

$$IN \rightarrow [[CONV \rightarrow ACT]^* N \rightarrow POOL]^* M \rightarrow \rightarrow [FC \rightarrow ACT]^* L, \quad (1)$$

где IN – входы СНС; $CONV$ – свёрточный слой; ACT – слой активации; $POOL$ – опциональный слой субдискретизации или пуллинга; FC – полносвязный слой. Параметры повторения принимают следующие значения: $0 < N < 3$, $M > 0$, $0 < L < 3$.

Слой пуллинга вычисляет одно значение матрицы на основе некоторой области исходного изображения, которая может иметь разные размеры, например, 2×2 .

В данном случае ИНС представляет собой стек из нескольких чередующихся слоев свёртки и активации, за которыми может следовать слой пуллинга. Затем идут полносвязные слои, последний из которых является выходом сети.

На рисунке 1 приведён пример модели, используемой для классификации. Она содержит пять свёрточных блоков для более точного выделения признаков, позволяющих определить отличия объектов. Размер ядра свёртки 3×3 для всех слоев. Каждый такой блок завершается операцией пуллинга, сокращающей размер матрицы признаков, которая будет передана на вход следующего блока. Окончательную классификацию выполняют полносвязные слои.

При описании свёрточных слоев обязательное задание ряда параметров свёртки, к которым относятся:

- количество фильтров, определяющее глубину слоя;
- размер ядра свёртки;
- шаг, с которым будет выполняться свёртка, в модели он $(1,1)$ по координатам (x, y) для всех слоев;

– параметр, определяющий, что изображение будет дополняться по краям нулями для сохранения исходного размера после прохождения свёрточных слоев.

Каждый фильтр свёрточного слоя нейронной сети отвечает за определение некоторого признака в изображении [3].

После определения последовательности слоев требуется выбрать нелинейную функцию активации для каждого из слоев, где это необходимо. В работе используется функция «выпрямитель», которая реализует пороговый переход в нуле и имеет вид:

$$f(x) = \max(0, x). \quad (2)$$

Преимуществом этой функции перед другими возможными функциями активации является то, что она легко вычисляется, не требуя выполнения ресурсоемких операций, что имеет существенное значение для обучения нейронной сети на больших объемах данных.

Для последнего полносвязного слоя сети, который отвечает за конечный результат классификации, может быть использована экспоненциальная функция «softmax».

$$f_j(x) = \frac{e^{x_j}}{\sum_{k=1}^K e^{x_k}}. \quad (3)$$

Данная функция позволяет определить вероятность отнесения объекта к каждому из K заданных классов, что приводит к получению результата классификации в виде вектора вероятностей размерности K .

Модель колоризации обычно имеет вид, подобный модели классификации, но с использованием только свёрточных преобразований признаков. За счёт способности нейронной сети к абстрагированию возможно использование цвета, отсутствовавшего на изображениях, использованных для обучения модели. Например, введение цвета менее насыщенного, чем на оригинальном.

Неотъемлемой частью данной модели является классификатор. При этом признаки, которые автоматически выделяются в ходе его работы, поступают на вход колоризатора. Признаки выделяются постепенно, от слоя к слою. Чтобы иметь возможность учитывать при колоризации все признаки, используются гиперколонки [4]. Гиперколонка представляет собой вектор выходов всех свёрточных слоев для каждого отдельного пикселя изображения, как показано на рисунке 2. Множество таких векторов для всех пикселей будет определять карту признаков всего изображения.

Пример чёрно-белого изображения и соответствующего результата работы такой модели колоризации представлен на рисунке 3.

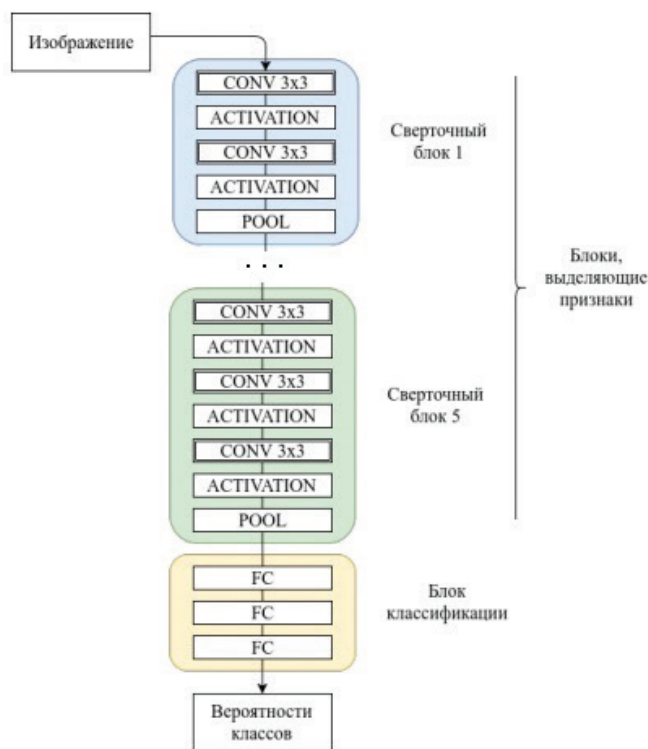


Рис. 1. Структура модели классификации на основе СНС
Fig. 1. Structure of classification model based on convolutional neural network

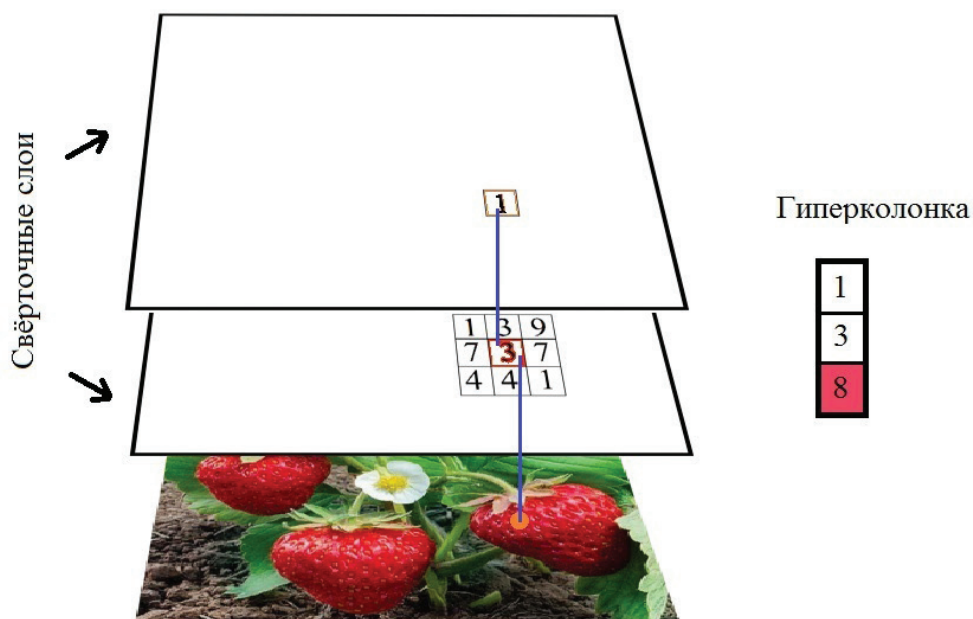


Рис. 2. Процесс извлечения гиперколонок
Fig. 2. Hypercolumns extraction process



Рис. 3. Пример работы модели колоризации на основе двух ИНС
Fig. 3. Sample of work result for colorization model based on two neural networks

Необходимо отметить, что существуют разные способы колоризации изображений, основанные на использовании сигналов в слоях СНС. Помимо использования специальной ИНС, преобразующей гиперколонки в цвета изображения, современным способом колоризации является остаточный кодировщик (Residual Encoder) [5]. Он основан на преобразовании выходных сигналов в слоях классифицирующей ИНС и не использует гиперколонки.

Алгоритм колоризации изображений, основанный на хеш-таблицах. Для преобразования гиперколонок в цвета изображения можно применять различные способы. Так, для хранения соответствия гиперколонок и цвета можно использовать стандартные структуры данных, например,

массивы или хеш-таблицы. В ходе исследований был разработан и реализован алгоритм колоризации изображений, основанный на СНС, которая выполняет классификацию изображений. Ее гиперколонки преобразуются в цвета без использования второй сети.

Последовательность шагов алгоритма:

1. Классифицировать обучающую чёрно-белую версию изображения с помощью СНС.

При этом необходимо сохранить выходные значения слоев СНС, кроме полносвязных и выходного слоя.

2. Рассчитать значения гиперколонок.

На данном этапе выполняется масштабирование выходных сигналов на слоях сети до размера входного сигнала.

3. Сохранить пары «гиперколонка – соответствующий цвет».

Если данная гиперколонка уже была сохранена, то к ней в соответствие ставится дополнительный цвет. Такую гиперколонку не рекомендуется в дальнейшем использовать для колоризации.

В качестве хранилища предлагается использовать хеш-таблицы. В роли хеш-функции может выступать некоторая простая функция от гиперколонки, например, взвешенная сумма её компонентов. Это позволяет получить выигрыш в скорости колоризации, по сравнению с использованием нейронной сети. По окончании обучения также можно построить упорядоченный список пар «гиперколонка – соответствующий цвет» на основе хеш-таблицы. Если предположить, что гиперколонки у каждого изображения уникальны, то размер требуемой памяти растёт линейно с ростом размера обучающей выборки (в отличие от нейронной сети, у которой количество нейронов фиксировано).

4. Повторять шаги 1–3 для всех изображений, используемых для обучения модели.

5. Классифицировать изображение с помощью той же нейронной сети, которая использовалась на первом шаге.

6. Для каждой полученной гиперколонки найти соответствие цвета в хеш-таблице. Если данной гиперколонки в ней нет, то найти наиболее близкую колонку и соответствующий ей цвет

методом бинарного поиска по упорядоченному списку.

В случае, если найденная колонка существенно отличается от искомой, принимается решение не раскрашивать соответствующий пиксель. Он остается исходного цвета.

7. Выполнить колоризацию изображения, используя найденные цвета.

8. Применить фильтр (например, медианный фильтр) для уменьшения шумов в полученном изображении.

Таким образом, данный алгоритм предполагает использование модели колоризации, содержащей классификатор на основе СНС, и колоризатор на основе хеш-таблицы.

Пример получившегося изображения представлен на рисунке 4. По сравнению с перечисленными ранее способами колоризации, есть потеря в равномерности определения цвета и зашумленности изображения, но полученный цвет является более насыщенным за счёт использования цветов обучающих изображений без изменений. Кроме того, имеется преимущество в скорости работы алгоритма благодаря использованию хеш-таблиц. Обработывались изображения размером 224x224 пикселей. Для улучшения чёткости может быть применено наложение полученного результата на исходное изображение.

Значения гиперколонок определяются структурой нейронной сети, использованной для классификации. Критериями для поиска опти-

Исходное изображение



Результат колоризации



Результат наложения

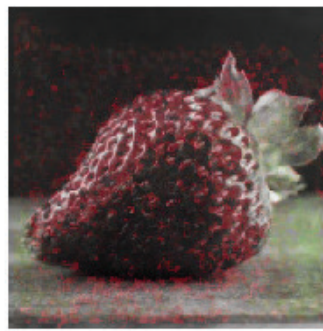


Рис. 4. Пример работы модели колоризации на основе ИНС и хеш-таблицы
Fig. 4. Sample of work result for colorization model based on neural network and hash table

мальной модели являлись площадь колоризации, корректность выбора цвета и скорость вычисления гиперколонок (не более 15 минут).

В результате была выбрана модель, содержащая СНС со следующими слоями: CONV 1x1 (15 ядер), ACTIVATION, CONV 3x3 (25 ядер), POOL 2x2, CONV 5x5 (4 ядер), CONV 5x5 (4 ядер), POOL 2x2, CONV 7x7 (4 ядер), CONV 15x15 (4 ядер), POOL 2x2, FC.

При верификации модели вычисляется значение потерь и точности выдаваемого моделью результата для каждого тестового изображения, и возвращается отношение успешно пройденных тестов к общему их количеству. По завершении колоризации, которая включает в себя денормализацию полученного изображения, выполняется его преобразование из числового массива в объект изображения. Разработанная модель и алгоритм реализованы в программном средстве, которое выполняет обучение ИНС и колоризацию графической информации.

Заключение. В работе предложены модель и алгоритм колоризации изображений, основанные на ИНС и хеш-таблицах.

Преимуществами разработанной модели являются более насыщенные цвета и более высокая скорость получения результата по сравнению с рассмотренными моделями. Для раскрашивания она использует только те значения цветов, которые выбирались на изображениях из обучающей выборки. К минусам можно отнести большее количество шумов. Пока предложенная модель лучше справляется с небольшим количеством классов и небольшой обучающей выборкой. Планируется её улучшение с целью устранения названных недостатков.

Модель и алгоритм колоризации изображений, описанные в данной статье, могут быть использованы для раскрашивания черно-белых фильмов, архивных фото, эскизов, для цветокоррекции цветных изображений и др.

Список литературы

1. Vieira, A. Introduction to deep learning business applications for developers / A.Vieira, B.Ribeiro. – New York: Apress, 2018. – 343 p.
2. Николенко, С. И. Глубокое обучение. Погружение в мир нейронных сетей / С. И. Николенко, А. А. Кадурин, Е. В. Архангельская. – СПб.: Питер, 2018. – 480 с.
3. An Intuitive Explanation of Convolutional Neural Networks [Electronic resource]. – Mode of access: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>. – Date of access: 20.04.2020.
4. Behnke, S. Hierarchical Neural Networks for Image Interpretation / S. Behnke. – Berlin: Springer, 2003. – 227 с.
5. Automatic Colorization [Electronic resource]. – Mode of access: <https://tinyclouds.org/colorize/>. – Date of access: 20.04.2020.

References

1. A. Vieira, B. Ribeiro. Introduction to deep learning business applications for developers. New York: Apress, 2018. 343 p.
2. S. Nikolenko, A. Kadurin, E. Arhangelskaya. Glubokoe obuchenie. Pogruzhenie v mir nejronnyh setej [Deep learning. Immersion in the world of neural networks]. SPb.: Piter, 2018. 480 p. (in Russian).
3. An Intuitive Explanation of Convolutional Neural Networks. Available at: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/> (accessed: 20.04.2020).
4. S. Behnke. Hierarchical Neural Networks for Image Interpretation. Berlin: Springer, 2003. 227 p.
5. Automatic Colorization. Available at: <https://tinyclouds.org/colorize/> (accessed: 20.04.2020).

Received: 25.04.2020

Поступила: 25.04.2020