Люди могут видеть широкий спектр цветов, что облегчает идентификацию объектов. Цветные изображения часто используются в различных областях для передачи значений и в творческих целях. Раскрашивание предполагает добавление цвета к изображениям в оттенках серого, что может быть выполнено с использованием различных методов, таких как подходы, ориентированные на пользователя, и подходы, основанные на обучении. Методы, основанные на обучении, становятся все более популярными, поскольку они требуют меньшего взаимодействия с человеком и их проще внедрять. Однако несбалансированное распределение по классам может привести к тому, что обучающие модели неправильно классифицируют наблюдения за классами меньшинств.

Проблема раскрашивания часто связана с распределением объектов, при этом несбалансированное распределение вызывает дисбаланс в процессе обучения. Это может привести к таким проблемам, как преобладание ненасыщенных цветовых компонентов в обучающих изображениях, что может повлиять на производительность насыщенных цветовых компонентов. Устранение дисбаланса функций важно для достижения желаемых результатов обучения.

Одним из способов решения этой проблемы является преобразование непрерывных значений цвета в дискретные цветовые классы и наоборот. Это может помочь предсказать распределение возможных цветов для каждого пикселя. В ходе исследования исследователи обнаружили, что преобладающе присутствовали 215 из 400 цветовых классов. Они уменьшили класс до 215, чтобы повысить точность классификации. Чтобы решить проблему дисбаланса классов, они определили веса классов, проанализировав истинный класс каждой партии во время обучения, присвоив более высокие веса редко появляющимся классам.

Исследователи также предложили метод селективной цветовой гармонизации объектов с помощью SAM для уточнения и полировки края. Они сформулировали задачу регрессии в виде задачи классификации, основанной на работе CIC. Они внедрили SAM для улучшения результатов за счет устранения цветовых искажений и обеспечения объектно-выборочной цветовой гармонизации в случаях сбоя модели.

В исследовании также была предложена новая метрика оценки цветовых изображений - отношение хроматических чисел (CNR), которая количественно определяет богатство цветовых классов в сгенерированных изображениях по сравнению с исходными изображениями. Исследование показало, что предложенный метод значительно превосходит существующие современные базовые показатели и дает приемлемые результаты.

## Раскрашивание изображений в основном можно разделить на две категории: раскрашивание под руководством пользователя и раскрашивание на основе обучения.

**A. Раскрашивание под руководством пользователя**

Раскрашивание под руководством пользователя можно дополнительно разделить на две категории: на основе каракулей и на основе примеров.

Раскрашивание на основе каракулей (не знаю как корректно перевести скрибл)

Методы раскрашивания на основе каракулей используют пользовательский ввод для заполнения отсутствующих или неполных участков изображения. Некоторые из методов, используемых при раскрашивании на основе каракулей, включают оптимизацию распространения цвета, комбинирование неитеративных методов с адаптивным выделением краев, введение смешивания цветов, эффективное распространение цвета в областях, непрерывных по рисунку, и областях, непрерывных по интенсивности, и включение структур U-Net.

Раскрашивание на основе примеров

Цветопередача на основе примеров сводит к минимуму усилия пользователя при передаче изображения в оттенках серого. Этот метод включает глобальную цветовую статистику, подходы на уровне сегментированных областей, методы на уровне суперпикселей и методы на уровне пикселей. Однако ручные показатели сходства могут приводить к ошибкам в сценариях со значительными различиями в интенсивности и содержании.

**B. Раскрашивание на основе обучения**

Раскрашивание на основе обучения - это метод машинного обучения, который автоматически применяет цвет к полутоновым или черно-белым фотографиям с использованием CNN, обученных на больших наборах данных. Основная проблема заключается в балансировке функций для сфокусированных объектов и фона.

Базовая раскрашивание на основе регрессии

Раскрашивание предполагает использование обычных CNN или специализированных архитектур, таких как InceptionNet, VGGNet, ResNet и DenseNet, для оценки цветовых каналов по изображениям в оттенках серого. Градиент вычисляется с использованием функции регрессионных потерь автоматизированными методами.

Раскрашивание на основе сегментации объектов

Различные модели раскрашивания сегментируют объекты внутри изображения, изучают назначение цветов по сегментам или объектам и назначают цвета сегментам, используя такие методы, как пространственные связи или глобальная цветовая согласованность. Эти модели были разработаны с использованием различных методов, включая семантическую сегментацию, состязательные модели с учетом границ и точечные аннотации.

Раскрашивание на основе GAN

Модели раскрашивания изображений GAN объединяют сети дискриминаторов и генераторов для получения реалистичных, эстетически приятных раскрашенных фотографий. В этих моделях используется семантическая информация, CapsNet, кодеры GAN и другие методы для улучшения результатов раскрашивания.

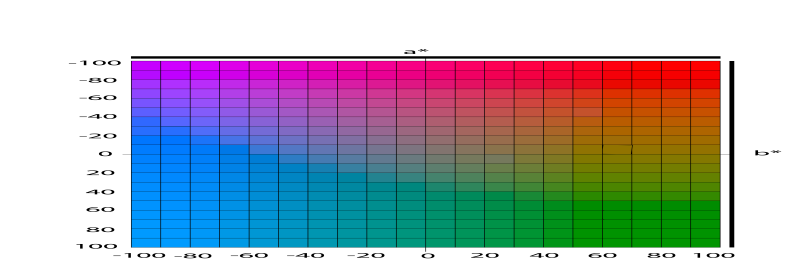
Балансировка функций для раскрашивания

Балансировка характеристик имеет решающее значение для методов раскрашивания. Чжан и др. предложили автоматическую раскраску с использованием CNN, классифицирующую интенсивность по заранее определенным уровням цвета и назначающую соответствующие цвета на основе классифицированных уровней класса. An и др. использовали модель CNN VGG-16 и метод перебалансировки цвета для решения проблем дисбаланса признаков. Ларссон и др. использовали несбалансированную потерю классификации, а Gain и др. предложили глубоко локализованную сеть для раскрашивания изображений.  
  
Алгоритм работы:

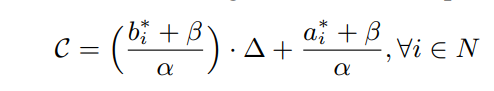
**Переводим изображение из rgb в L\*a\*b**

В пространстве La\*b\* (LAB) L обозначает яркость изображения, при этом значения интенсивности находятся в диапазоне [0, 100]. При увеличении L цвета становятся ярче. Каналы a\* и b\* соответственно представляют соотношение красно-зеленых и желто-синих тонов изображения, причем красно-желтый представлен положительным значением, а зелено-синий - значительным отрицательным значением, часто находящимся в диапазоне [-128, 127].

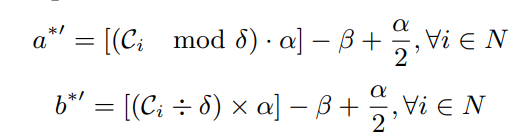
При фиксированном L каналы а и b не различимы для человека физически, что позволяет перевести задачу по раскраске в задачу классификации по парам a b с непрерывными значениями.



Каждую точку на этом поле можно отнести к определённому классу таким уравнением



Что бы перевести из класса, предсказанного ИИ, нужно воспользоваться формулами:



Согласно приведенным выше уравнениям, максимальная потеря для каждого

значения a\* или b\* равна . Более высокое значение α уменьшает

количество классов, но приводит к потерям в представлении, поскольку

большой непрерывный диапазон преобразуется в один класс. Однако

решить проблему с меньшим классом легко. Меньшее

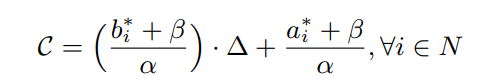
значение α увеличивает количество классов. В задаче раскрашивания

большее количество классов делает прогноз менее точным.

Важно настроить количество уровней класса для цветового пространства a\*b\*

таким образом, чтобы модифицированный a\*b\* мог описывать природу цвета изображения.

**Уменьшение цветового класса на основе практичного внешнего вида**

Значения цвета a\* и b\* постоянны в пределах [-128, 127] в цветовом канале a\*b\*. Но на практике диапазон находится в пределах [-100, 99]. Сначала мы преобразуем непрерывный [-100, 99] распределил цветовой канал a\*b\* по одной плоскости из 400 цветовых классов, взяв α = 10, β = 100 и ∆ = 20 в 

Затем формируется двумерная сетка ячеек в виде одноплоскостного массива размером 20 × 20, где горизонтальные оси указывают a\*, а вертикальные оси указывают информацию о цвете b\*.

Исследователи заметили, что цвета меняются очень постепенно, и ближайшие классы очень похожи, поэтому они сжали все изображения до 56х56. Это привело к тому, что количество классов уменьшилось, а очень редкие объекты стали встречаться реже, чтобы это избежать, изображения можно сделать меньше.

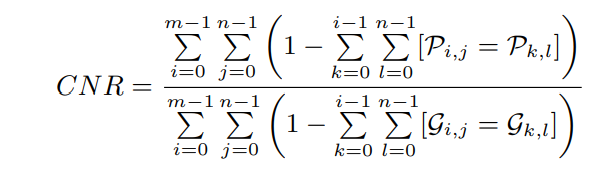
//здесь идет ещё тонна информации для построения нейросети, которые я пропущу за //ненадобностью для понимания работы самого алгоритма

Псевдокод

1. Подается изображение в LAB формате
2. С помощью модели SAM выделяются значимые блоки (S)
3. For s in S:
   1. If число пикселей в s > Ψ:
      1. Считаем моду по каналу a\*
      2. For каждый пиксель в области:
         1. If модуль разности пикселя с модой больше дельты:
            1. Заменить значения пикселя на моду

//аналогичное проделываем и для b\* канала

Для обучения нейросети исследователи ввели новую метрику



Где P – сгенерированные нейросетью пиксели, а G - естественный цвет.

Выражение в скобках возвращает 1 в случае равенства и 0 в обратном случае.

После тестов у исследователей получились такие коэффициенты:

