

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ И КОМПЬЮТЕРНОГО ЗРЕНИЯ В ОЦЕНИВАНИИ ДЕТСКИХ РИСУНКОВ

Д.Ю. Немцова, А.А. Серов

ФГБОУ ВО «Тверской государственный университет», Тверь, Россия

Рассматривается применение методов машинного обучения и компьютерного зрения для задачи оценивания качества рисунков детей младшего школьного возраста. Представлены технологии машинного обучения для анализа качества стандартных субъективных критериев оценивания и возможные методы оценивания рисунков с применением технологий компьютерного зрения. Исследованы связи между субъективными экспертными оценками детских рисунков и некоторым машинным аналогами таких оценок на основе фрактальной размерности. Разработана соответствующая регрессионная модель. Объясняются причинно-следственных связей между данными характеристиками изображений с применением байесовских сетей.

Ключевые слова: критерии оценивания детских рисунков, компьютерное зрение, фрактальная размерность, анализ надежности теста, альфа Кронбаха, омега-статистика, тепловая карта корреляционной матрицы, регрессионная модель, байесовские сети, сверточные нейронные сети.

В настоящее время технологии машинного обучения и искусственный интеллект внедряются во все сферы современного общества, в том числе и школьного образования. Данный процесс находит свое применение и в современном художественном образовании: генерация изображений, преобразование изображений. К данной сфере достаточно близки технологии виртуальной (VR) и дополненной (AR) реальностей: виртуальный музей, картинная галерея, погружение в мир картины известного художника и др.

В настоящей работе предлагается использование методов машинного обучения и компьютерного зрения на основе специальных статистик в задаче оценивания детских рисунков:

1. Выполнен анализ дифференцирующих способностей субъективных стандартных экспертных оценок в целом и каждого критерия в отдельности;
2. Исследованы связи фрактальной размерности изображения с субъективными критериями и с суммарными оценками.
3. Приведены примеры вычислений индекса структурного сходства цифровых изображений двух детских рисунков, их разности и расстояния между ними по хешам.
4. Разработана примерная модель оценки качества рисунков с применением сверточных нейронных сетей.

В данной работе проведён анализ живописных тематических детских рисунков. Их создали дети в возрасте от 4 до 15 лет с различным уровнем подготовки. Произведения выполнены в свободной технике с использованием различных материалов.

Вопрос оценки детского изобразительного творчества достаточно сложный: понятия правильности или соответствия стандартам достаточно условны. Для того, чтобы свести к минимуму субъективность, необходимо привлечь несколько специалистов данной сферы и использовать специальные критерии оценивания. Под критериями оценивания детских рисунков понимают наличие у обучающихся основ изобразительной грамоты по следующим

параметрам: 1) композиция: гармоничное расположение объектов на листе, их единство и пропорциональность; 2) форма: передача характерных особенностей объектов; 3) пространство: наличие перспективы; 4) цвет: передача действительной окраски предметов. Относительно композиции – наличие гармонии, единство и цельность пятен; 5) тон: насыщенность цвета.

Для оценки живописных тематических работ используют критерии, которые могут корректироваться в зависимости от конкретных условий, целей и задач. В своей работе при оценивании рисунков мы использовали следующие критерии: 1) наличие цветовой гармонии и тонального решения (x1); 2) композиционное решение (x2); 3) художественная выразительность (x3); 4) творческий подход, оригинальность авторского решения (x4); 5) раскрытие темы (x5); 6) владение художественными материалами (x6); 7) аккуратность (x7); 8) передача пропорций (x8); 9) передача пространства (x9); 10) законченность работы (x10) [1].

На наш взгляд, по этим критериям можно с максимальной точностью определить, насколько качественны данные работы в творческом отношении. Система оценивания по каждому критерию – десятибалльная, где 1 балл – плохое качество и 10 баллов – отличное. При оценивании живописных работ учитывался возраст обучающихся. Исходные данные представлены в форме таблицы, в столбцах которой находятся экспертные оценки 100 различных детских рисунков (по всем критериям) и вычисленные оценки фрактальной размерности Минковского-Булиганда для каждого рисунка. Компьютерные характеристики были вычислены в среде Python 3 на основе методов, предложенных в [3].

В данном исследовании основное место отведено фрактальной размерности изображения и ее связям с субъективными оценками детских работ. Понятие фрактальной размерности изображения связано с применением сложной математической теории и не имеет простого описания. Говоря о фрактальной размерности, можно только определить, что она характеризует «сложность», «самоподобие» или специальную «симметрию» изображения. Существует ряд методов и подходов к определению фрактальной размерности изображения, среди которых чаще используется размерность Минковского-Булиганда. Фрактальная размерность основана на связи изменений в элементе изображения с изменениями в его масштабе. Ее числовое значение равно угловому коэффициенту некоторой линии регрессии для некоторых специальных данных, вычисленных по изображению.

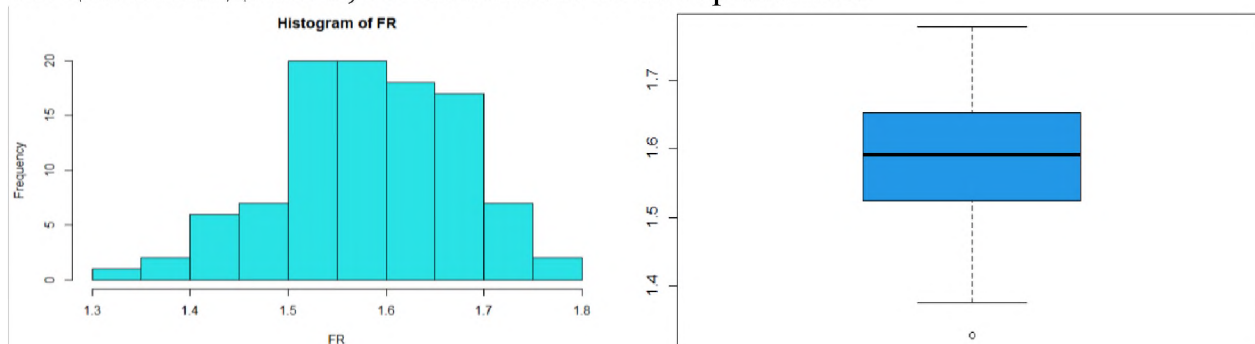


Рис. 1. Гистограмма частот (слева) и диаграмма размахов значений фрактальной размерности (справа) детских рисунков для данных в исследовании

Были проведены исследования взаимосвязей между всеми исходными показателями оценивания детских рисунков: субъективными и указанными компьютерными оценками (фрактальной размерностью). С этой целью были вычислены ранговые корреляции Спирмена для всех переменных. Результаты корреляционного анализа представлены на рис. 2.

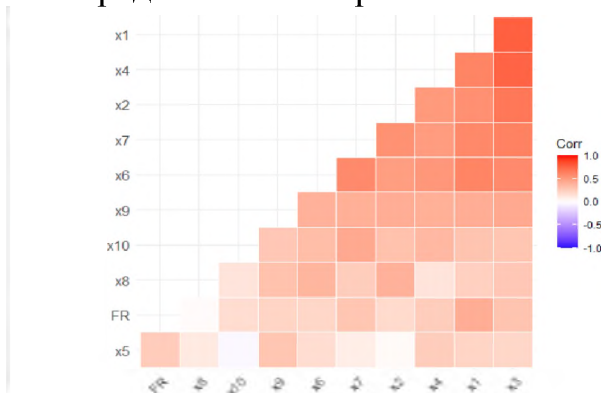


Рис. 2. Тепловая карта корреляционной матрицы для всех исходных показателей оценивания детских рисунков.

Попробуем объяснить некоторые корреляционные связи между показателями (переменными): 1) художественная выразительность (x3) данных работ зависит от средств живописи: цвета, мазков, пятен, цветового и светового контраста, колорита (x1); 2) в творческом подходе (x4) проявляется художественная выразительность (x3); 3) цветовое решение (x1) повторяется в различных элементах изображения, что может быть связано с фрактальной размерностью (FR).

На основе полученных эмпирических данных был выполнен анализ пригодности (надежности теста) применяемых субъективных критериев оценивания рисунков и сравнительный анализ этих критериев. Визуализация результатов данных вычислений представлена на рис. 3.



Рис. 3. Диаграмма для сравнения дифференцирующих способностей субъективных критериев x1-x10 для оценивания детских рисунков. Числовое значение альфы Кронбаха (0.86) всего теста обозначено красной линией.

В статистической среде R с применением пакета psych были получены следующие результаты: альфа Кронбаха всего теста: 0.86 с 95% доверительным интервалом (0.80, 0.89); G6 Гуттмана: 0.89, омега иерархическое: 0.77; омега общее: 0.91. Можно сделать вывод, что качество всего теста и отдельных критериев очень высокое: все критерии оценивают одно и то же свойство детей – умение рисовать.

Для изучения влияний переменных друг на друга были построены байесовские сети для всех исходных показателей (рис. 4). Вычисления выполнены в статистической среде R с применением пакета bnlearn.

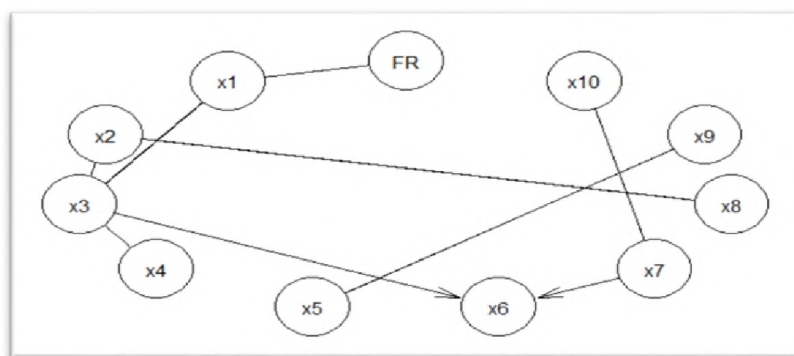


Рис. 4. Байесовская сеть взаимного влияния для субъективных критериев оценивания детских рисунков (x1-x10) и фрактальной размерности (FR).

Обнаружена связь фрактальной размерности рисунков с критерием x1 и др.

Построим математическую модель, связывающую объективную оценку рисунка с субъективной, т.е. выполним регрессионный анализ для переменных: фрактальная размерность (FR, независимая переменная) и субъективная суммарная оценка детских рисунков по стандартным тестам (TS, зависимая переменная). Коэффициент корреляции Пирсона для этих переменных равен 0.404 с 95-процентным доверительным интервалом (0.256, 1.000), значимость (p-value) равна 0.0002. Таким образом, обнаружена статистически значимая положительная корреляция между фрактальной размерностью цифрового изображения детского рисунка и его субъективной экспертной оценкой. Данный результат наших исследований можно считать достаточно интересным и существенным: найдена одна из автоматически вычисляемых объективных оценок детских рисунков, которая достаточно сильно и положительно коррелирует с суммарной субъективной экспертной оценкой этих рисунков. Построим соответствующую регрессионную модель в среде R. Полученная модель статистически значима со всеми статистически значимыми коэффициентами:

$$\text{Суммарная оценка} = 28.915 * \text{Фрактальная размерность} + 45.825.$$

Скорректированный коэффициент детерминации R^2 равен 0.155. Следовательно, в данной регрессионной модели объясняется почти 16% изменчивости субъективной оценки детских рисунков. Визуализации этой модели представлены на рис. 6.

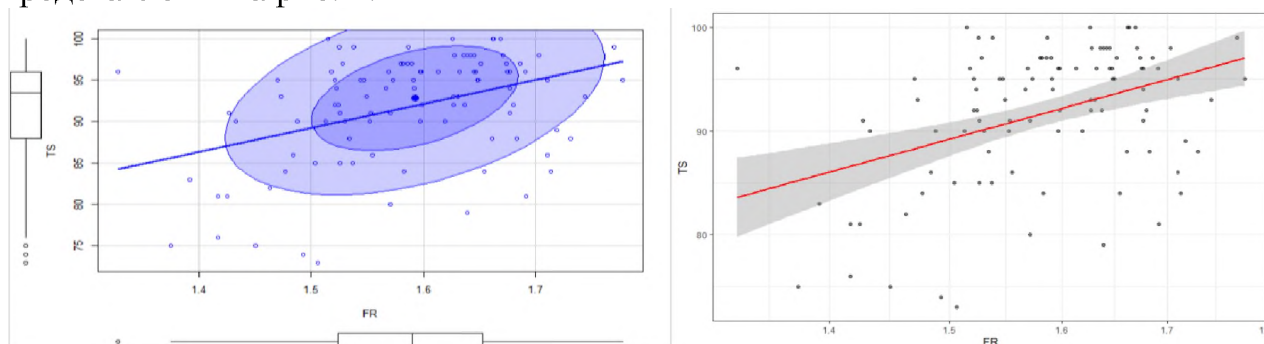


Рис. 6. Визуализации простой регрессионной модели зависимости субъективной суммарной оценки (TS) от фрактальной размерности (FR). Слева: с эллипсами концентрации точек (50% и 95%) и диаграммами размахов; справа: с 95% доверительной областью линии регрессии

Методы определения качества изображений (рисунков) и субъективно, и с применением технологий машинного обучения могут быть разделены на

референтные (сравнение данного изображения с изображением-образцом) и безреферентные (без использования образца). Все перечисленные выше методы являются безреферентными. В машинном обучении и в компьютерном зрении существуют и различные референтные методы определения качества изображений: на основе индекса структурного сходства двух изображений SSIM [4], с применением хешей изображений [6], разности изображений и др. Приведем примеры вычислений для некоторых из указанных методов сравнения двух изображений. Вычислим индекс структурного сходства для двух детских рисунков, представленных на рис. 7 (слева).



Рис. 7. Пример изображений двух детских рисунков для вычисления индекса структурного сходства SSIM и их разность (справа)

На разности изображений одинаковые фрагменты рисунков выделены черным цветом. Индекс структурного сходства SSIM данных изображений равен 0.15. Вычислим хеши данных рисунков и найдем расстояние Хемминга между рисунками по хешам [6]. Хеш первого изображения: e7ef3f7e60e09c9e; хеш второго изображения: 0016243c20ffffff. Расстояние Хемминга между рисунками по хешам равно 31.

С научной и методической точек зрения важное значение имеет создание в ближайшем будущем инструментального средства на основе технологий глубокого машинного обучения и компьютерного зрения для автоматического оценивания детского рисунка. Алгоритм применения подобного инструмента мог бы быть примерно следующим:

1. Сканирование рисунка с помощью смартфона.
2. Вычисление оценки сканированного рисунка с применением предварительно созданной специальной компьютерной модели (приложения) в сети Интернет с анализом изображения в форме текста и др.

Считаем, что первые шаги к созданию подобных средств оценивания в настоящее время уже сделаны. Методы машинного обучения, предназначенные для таких целей, могут быть основаны как на отдельных характеристиках изображения, так и на более сложных алгоритмах с применением сверточных нейронных сетей, сиамских сетей, индекса структурного сходства SSIM [4], индекса BRISQUE [5], хешей изображения [6] и др. Полагаем, что достаточно легко могут быть разработаны подобные инструменты оценивания детского рисунка не в целом, а по отдельным критериям, например, по композиции, цветовому решению и др. В нашем исследовании в рамках вычислительного эксперимента была разработана примерная модель сверточной нейронной сети с применением модуля керас [2] для классификации детских рисунков (по их цифровым изображениям). Все рисунки были разделены экспертом на два класса: сравнительно удачные и неудачные. Все изображения разбили на

обучающую и тестовую выборки в пропорции 4:1 в каждом классе изображений. В результате обучения нейронной сети была получена модель с точностью классификации на тестовой выборке около 70%. Данный результат является достаточно близким к удовлетворительному. Для повышения точности этой модели требуется разработка специальных баз изображений детских рисунков с различным качеством. Аналогичные сложности возникают и для вычислений индекса качества BRISQUE: нужна специальная база изображений детских рисунков с усредненными оценками нескольких экспертов. Эти подходы к оцениванию детских рисунков нуждаются в дальнейших исследованиях. Полагаем, что их результаты будут сильно зависеть и от вида художественного творчества детей.

СПИСОК ЛИТЕРАТУРЫ

1. Лыкова Е.С. Критерии оценки детского рисунка. URL: <https://clck.ru/aeEwj> (дата обращения: 28.12.2021).
2. Шолле Ф. Глубокое обучение на Python. / пер. с англ. А. Киселева. СПб.: Питер, 2020. 400 с.
3. Fractal dimension computing. URL: <https://clck.ru/aePqM> (дата обращения: 21.12.2021).
4. Scikit-image. URL: <https://scikit-image.org/> (дата обращения: 21.12.2021).
5. Image-quality 1.2.7. URL: <https://pypi.org/project/image-quality/> (дата обращения: 21.12.2021).
6. Image Hashing library 4.2.0. URL: <https://pypi.org/project/ImageHash/> (дата обращения: 21.12.2021).

Об авторах:

Немцова Дарья Юрьевна – студентка 3 курса направления 44.03.01 Педагогическое образование, профиль «Изобразительное искусство», ФГБОУ ВО «Тверской государственный университет» (Россия, 170100, г. Тверь, Желябова, 33); e-mail: dynemtsova@edu.tversu.ru

Серов Анатолий Александрович – кандидат физико-математических наук, доцент кафедры математического и естественнонаучного образования, ФГБОУ ВО «Тверской государственный университет» (Россия, 170000, г. Тверь, Желябова, 33); e-mail: Serov.AA@tversu.ru

METHODS OF MACHINE LEARNING AND COMPUTER VISION IN ASSESSING CHILDREN'S DRAWINGS

D.Yu. Nemtsova, A.A. Serov

Tver State University, Tver, Russia

The application of machine learning and computer vision methods for the problem of assessing the quality of drawings of primary school children is considered. Machine learning technologies for analyzing the quality of standard subjective assessment criteria and possible methods for assessing drawings using computer vision technologies are presented. The connections between subjective expert assessments of children's drawings and some machine analogs of such assessments based on fractal dimension are investigated. A corresponding regression model has been developed. An attempt is made to explain the causal relationships between these characteristics of images using Bayesian networks.

Keywords: *criteria for evaluating children's drawings, computer vision, fractal dimension, test reliability analysis, Cronbach's alpha, omega statistics, heat map of the correlation matrix, regression model, Bayesian networks, convolutional neural networks.*

About the authors:

Nemtsova Darya Yuryevna – student direction 44.03.01 Pedagogical education, profile «Fine Arts», Tver State University (Zhelyabova 33, Tver 170000, Russia), e-mail: dynemtsova@edu.tversu.ru

Serov Anatoly Alexandrovich – Candidate of Physical and Mathematical Sciences, Associate Professor of Mathematics and Science Education Tver State University (Zhelyabova 33, Tver 170000, Russia); e-mail: Serov.AA@tversu.ru