

<https://doi.org/10.31891/2307-5732-2025-349-56>

УДК 004.383.5:004.032.26

ТЕГЛІВЕЦЬ ОЛЕСЬ

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0003-3880-2159>

e-mail: oles.v.tehlivets@lpnu.ua

ТЕГЛІВЕЦЬ ЮЛІЯ

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-7221-7554>

e-mail: yuliia.v.tehlivets@lpnu.ua

ОЦІНКА ЯКОСТІ ЗОБРАЖЕННЯ ЗА ДОПОМОГОЮ СУБ'ЄКТИВНОГО МЕТОДУ ОЦІНЮВАННЯ З ОДНИМ СТИМУЛОМ

У цій статті представлено дослідження суб'єктивного методу оцінювання зображень технології глибокого навчання, зокрема одного із його різновидів – методу з одним стимулом. Важливість суб'єктивної оцінки полягає в тому, що її використовують для визначення, як зображення виглядає для реального користувача. В основі цього підходу лежить психофізика зору, яка аналізує, як різні візуальні фактори впливають на сприйняття. Мета дослідження – проаналізувати критерії, за якими користувачі оцінюють якість зображень за допомогою суб'єктивних методів оцінки якості зображень; з'ясувати, як різні групи користувачів оцінюють якість зображення у реальних умовах за допомогою методу з одним стимулом виявити переваги та недоліки суб'єктивного методу оцінювання для моделі з одним стимулом. Важливим етапом дослідження було його проведення в реальних умовах без застосування спеціального обладнання. Для кращої репрезентативності вибірки було залучено різні групи користувачів з нормальним або скоригованим до нормального зором, щоб отримати різносторонні оцінки. Усіх респондентів було поділено на 4 групи: група N1 – респонденти, що належать до мистецької галузі, група N2 – респонденти – професійні фотографи та відеографи, група N3 – звичайні користувачі одного віку, група N4 – звичайні користувачі різного віку. Загалом взяло участь 60 респондентів. Для дослідження було обрано еталонне зображення з колекції бази даних TID2013 і використано шкалу MOS. Результати дослідження показують, що група N2 виявилась найвимогливішою до оцінювання, що відповідає професійному рівню респондентів, а група N4 – найменш критичною. Чіткість зображення більше важлива для професіоналів, але менш важлива для звичайних користувачів. Професіонали більше звертали увагу на технічні деталі, звичайні користувачі орієнтувалися на загальне враження та естетику зображення. Дослідження дозволило виявити позитивні та негативні сторони суб'єктивного методу оцінювання з одним стимулом. Серед переваг – висока релевантність: оцінка естетичних та емоційних аспектів, виявлення прихованих дефектів, диференціація за групами тощо. Серед недоліків можна відзначити часозатратність, ресурсозатратність, індивідуальність сприйняття, технічне забезпечення тощо. Огляд основних концепцій і методів демонструє необхідність інтеграції суб'єктивних оцінок із технологіями глибокого навчання для покращення оцінки якості зображень. Це сприятиме створенню адаптивних моделей, які краще відповідають людському сприйняттю.

Ключові слова: оцінка якості зображення, глибока нейронна мережа, навчання нейронної мережі, база даних TID2013, суб'єктивний метод оцінювання зображення.

TEHLIVETS OLES

TEHLIVETS YULIYA

Lviv Polytechnic National University

IMAGE QUALITY ASSESSMENT USING A SUBJECTIVE EVALUATION METHOD WITH A SINGLE STIMULUS

This article presents a research of the subjective method of evaluating images of deep learning technology, in particular, one of its varieties - the single-stimulus method. The importance of subjective evaluation is that it is used to determine how an image looks to a real user. This approach is based on the psychophysics of vision; which analyses how various visual factors affect perception. The purpose of the study is to analyse the criteria by which users evaluate image quality using subjective image quality assessment methods; to find out how different groups of users evaluate image quality in real conditions using a single-stimulus method; to identify the advantages and disadvantages of the subjective assessment method for the single-stimulus model. An important stage of the study was to conduct it in real conditions without the use of special equipment. For better representativeness of the sample, different groups of users with normal or corrected-to-normal vision were involved to obtain diverse evaluations. All respondents were divided into 4 groups: group N1 - respondents belonging to the artistic industry, group N2 - respondents - professional photographers and videographers, group N3 - ordinary users of the same age, group N4 - ordinary users of different ages. A total of 60 respondents took part. A reference image from the TID2013 database collection was selected for the study and the MOS scale was used. The results of the study show that group N2 was the most demanding to evaluate, which corresponds to the professional level of the respondents, and group N4 was the least critical. Image clarity is more important for professionals, but less important for ordinary users. Professionals paid more attention to technical details, while ordinary users focused on the overall impression and aesthetics of the image. The study revealed the positive and negative aspects of the subjective single-stimulus evaluation method. The advantages include high relevance: assessment of aesthetic and emotional aspects, detection of hidden defects, differentiation by groups, etc. Among the disadvantages are time-consuming, resource-consuming, individual perception, technical support, etc. The review of basic concepts and methods demonstrates the need to integrate subjective assessments with deep learning technologies to improve image quality assessment. This will help to create adaptive models that better match human perception.

Keywords: image quality assessment, deep neural network, neural network training, TID2013 database, subjective method of image evaluation.

Постановка проблеми у загальному вигляді

Із розвитком цифрових технологій якість зображень стала важливим фактором у багатьох галузях, таких як цифрова фотографія, електронна комерція, медичні дослідження, відеоігри, AR/VR-

технології тощо. Цифрові зображення швидко проникають у наше повсякденне життя завдяки великому обсягу інформації у вигляді візуальних сигналів. Ці зображення часто проходять через кілька етапів оброблення, перш ніж потрапляють до кінцевого користувача. Користувачі очікують отримувати візуальний контент високої якості, незалежно від пристрою чи платформи. Відтак зростання попиту на високоякісні зображення в умовах обмежених ресурсів зумовлює необхідність покращення методів оцінки та підвищення якості зображень.

Глибоке навчання значно змінило підходи до оброблення зображень, дозволяючи автоматизувати покращення якості та адаптувати результати до потреб кінцевих користувачів. Такі нейронні мережі, як GAN (Generative Adversarial Networks) і CNN (Convolutional Neural Networks), стали потужним інструментом для покращення якості зображень. Так, GAN використовують для генерації високоякісних зображень, які важко відрізнити від оригінальних, CNN використовують для завдань суперрезолюції, де низькоякісне зображення перетворюється на зображення високої чіткості. Проте більшість цих моделей навчаються на об'єктивних метриках (PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity Index), LPIPS (Learned Perceptual Image Patch Similarity)), що може обмежувати їхню здатність створювати зображення, які відповідають людським очікуванням. Основна проблема полягає в тому, що об'єктивні метрики часто не враховують такі нюанси людського сприйняття, як естетика, текстура чи природність кольорів [1]. Наприклад, два зображення можуть мати однакові значення PSNR, але глядачі їх сприйматимуть по-різному через відмінності в деталях [2]. Суб'єктивні оцінки є важливими, оскільки вони враховують індивідуальне сприйняття деталей зображення, таких як кольорова гама, текстура, розмитість зображення, наявність артефактів, зернистості тощо. В основі цього підходу лежить психофізика зору, яка аналізує, як різні візуальні фактори впливають на сприйняття. Ці фактори можуть бути дуже важливими у таких галузях, як реклама, медична візуалізація, фото- та кіноіндустрія тощо, де загальне враження від зображення є ключовим для досягнення бажаного ефекту [2].

Аналіз досліджень та публікацій

Глибокі нейронні мережі стали основою багатьох сучасних підходів щодо оброблення зображень завдяки здатності нейромереж автоматично виявляти і виправляти дефекти оброблення зображень. Навчання згорткових нейронних мереж для опрацювання зображень зазвичай потребує значних обчислювальних ресурсів, використання графічних процесорів GPU [3].

Об'єктивні метрики, розроблені для кількісного оцінювання якості зображень на основі математичних обчислень, забезпечують швидку оцінку, однак часто не відповідають тому, як людина сприймає реальну якість зображення [1,4,5]. Це обумовлено їхньою неспроможністю врахувати всі аспекти людського зору, що призводить до суб'єктивності результатів. Наприклад, зображення з високим значенням PSNR може містити спотворення, які залишаються непомітними для метрики, але є очевидними для людського ока. Тобто користувач може відзначити, що зображення виглядає неприродно або має нереалістичні кольори, навіть якщо метрика показує високий рівень якості. [1]. Саме суб'єктивний метод оцінювання зображення дозволяє зрозуміти, як кінцевий користувач сприймає зображення, що є основним аспектом у багатьох застосуваннях [6-10].

Оскільки традиційні об'єктивні метрики не завжди відповідають реальному сприйняттю якості зображень, інтеграція суб'єктивних методів оцінки з глибоким навчанням є необхідною для створення більш точних моделей. Наприклад, модель NIMA (Neural Image Assessment), що навчається на суб'єктивних оцінках якості, використовуючи CNN (Convolutional Neural Networks), дозволяє прогнозувати, як користувачі оцінять зображення [11]. У додатках для редагування фотографій суб'єктивні оцінки допомагають оптимізувати фільтри так, щоб результати виглядали привабливими для більшості користувачів. Деякі науковці для додатків, в яких зображення в кінцевому підсумку буде переглядати людина, суб'єктивне оцінювання вважають єдиним правильним методом щодо визначення якості візуального зображення. Хоча й зазначають, що на практиці суб'єктивне оцінювання зазвичай занадто незручне, трудомістке і вартісне [1, 10-13].

Суб'єктивна оцінка якості зображень передбачає кілька основних критеріїв, що були запропоновані та розроблені в численних дослідженнях. Наприклад, було розглянуто концепцію «видимих артефактів», таких як шум, розмиття, різні спотворення та їхній вплив на сприйняття зображень. Також запропоновано моделі, які враховують психологічні аспекти сприйняття відео та зображень, вводячи поняття людського зору та якості зору. [14-18]. Наприклад, показники якості зображення/відео оцінюють у контексті стиснення зображення/відео, передавання та наявності водяних знаків. Відтак основні принципи суб'єктивної оцінки охоплюють такі параметри, як чіткість, природність кольорів, видимість артефактів.

Складність автоматизованого оцінювання якості зображення зумовлена особливостями зорової системи людини, яка сприймає як дрібні деталі зображення, так і все зображення в цілому. Оскільки людина часто є кінцевим споживачем візуальних продуктів (відео, зображення), суб'єктивна оцінка є найбільш підходить для отримання кількісного значення якості зображення в певній шкалі значень. Однак така оцінка не може бути безпосередньо використана в автоматизованій обробці зображень [19-20].

Поєднання суб'єктивних методів із технологіями глибокого навчання дозволяє створювати більш персоналізовані рішення, які враховують не лише технічні, але й естетичні аспекти зображень. Це

актуально для галузей, де сприйняття якості має вирішальне значення (наприклад, у кіноіндустрії, вебдизайні, електронній комерції) [21-25].

Формулювання цілей статті

Проаналізувати критерії, за якими користувачі оцінюють якість зображень за допомогою суб'єктивних методів оцінки якості зображень; з'ясувати, як різні групи користувачів оцінюють якість зображення у реальних умовах за допомогою методу з одним стимулом виявити переваги та недоліки суб'єктивного методу оцінювання для моделі з одним стимулом.

Виклад основного матеріалу

Метод з одним стимульним зображенням (Single-stimulus methods), який ми обрали для нашого дослідження, передбачає показ зображення на короткий і фіксований час, а потім просять респондентів оцінити його. Зображення було обрано з колекції бази даних TID2013, розробленої для оцінювання та порівняння показників візуальної якості зображень та експертних числових оцінок. Важливою особливістю цієї бази зображень є велика кількість типів спотворень, яка охоплює різні умови отримання зображень, різні рівні освітленості, контрасту, різні типи шумів, втратні стискання зображень тощо. Для кожного зображення наведено середню експертну оцінку його якості, що забезпечує можливість їх використання для навчання нейронних мереж. Крім того, цю базу зображень можна розширити додаванням нових рівнів спотворень [3, 20, 26].

Враховуючи можливі психологічні фактори, зокрема, щоб уникнути ефекту втоми, ми обрали одне зображення (рис. 1). Це зумовлено тим, що оцінюючи багато варіантів респонденти можуть давати упереджені або випадкові оцінки, щоб швидше завершити тестування.



Рис.1. Приклад еталонного зображення для оцінювання

Зображення оцінювали наївні спостерігачі і професіонали з нормальним або скоригованим до нормального зором. Загалом взяло участь 60 респондентів., з них 38 жінок і 22 чоловіки. Вік респондентів варіювався від 18 до 60 років.

Для кращої репрезентативності вибірки було залучено різні групи користувачів, щоб отримати різносторонні оцінки. Усіх респондентів було поділено на 4 групи. Кожна група налічувала однакову кількість респондентів – 15.

Група N1 – респонденти, що належать до мистецької галузі. Їх склали студенти таких напрямків, як дизайн та архітектура.

Група N2 – респонденти – професійні фотографи та відеографи.

Група N3 – звичайні користувачі одного віку. Їх склали студенти комп'ютерної галузі.

Група N4 – звичайні користувачі різного віку. Остання група – це так звані «наївні» респонденти, в тому сенсі, що вони не вивчали і не працювали з якістю зображень або не працюють у сферах, що пов'язані із ними. Безпосередньо перед тестуванням було зібрано базову інформацію щодо цієї групи, яка включає в себе такі моменти: звички до фотографування, власне фотообладнання, вміння обробляти зображення в графічних редакторах. Для експерименту щодо цієї групи обрано респондентів, які практично не мають навичок з оброблення зображень і використання професійної техніки для фотографування.

Учасники оцінювали якість зображення в умовах, наближених до реального використання, на екранах різних пристроїв (смартфон, комп'ютер). Як зазначають, якісні експерименти рідко проводять у контрольованих умовах. Це пов'язано з тим, що в реальних умовах зображення сприймають з різної відстані на екранах з різною роздільною здатністю. Тому дані є більш репрезентативними для реальних умов [21]. Для усіх груп респондентів було обрано однаковий час у першій половині дня, щоб усі учасники експерименту оцінювали зображення при достатній кількості світла.

Середня тривалість експерименту – 30 хвилини. Цей час включав інструкції для респондентів та безпосередньо візуальне тестування. Спостерігачі також мали можливість зробити перерву, якщо відчували, що вона їм потрібна.

Для оцінювання використано шкалу MOS (Mean Opinion Score), яка є основним показником у таких методах. Респонденти оцінювали зображення за шкалою від 1 до 5, де 1 означає «дуже погано», а 5 – «дуже добре». Середню оцінку думки (MOS) було розраховано на основі загальної. Цей метод використовують у стандартах ITU-T для оцінки якості відео та зображень [27, 28]. Метод з одним

стимулом вважають доволі ефективним, оскільки він вимагає лише n+1 спробу для оцінки n умов (одна додаткова спроба для еталонного зображення).

Група N1 (рис. 2), зважаючи на профіль групи, продемонструвала орієнтацію на естетику, гармонію кольорів і композицію. Такий параметр, як природність кольорів (4.27), було високо оцінено, але без максимальних значень, що вказує на збалансованість сприйняття кольорів на зображенні. Високо оцінено загальну якість (4.13), чіткість (4.2), загальне враження (4.13), естетичність (4.07) і загальний рівень шуму (зернистості) (4). Порівняно низький бал отримали баланс білого (3.73), контрастність (3.73) і яскравість (3.6), що може свідчити про те, що ці критерії менш важливі для представників мистецької галузі, або про недоліки, які потребують доопрацювання. Найнижчим показником оцінено яскравість (3.6).

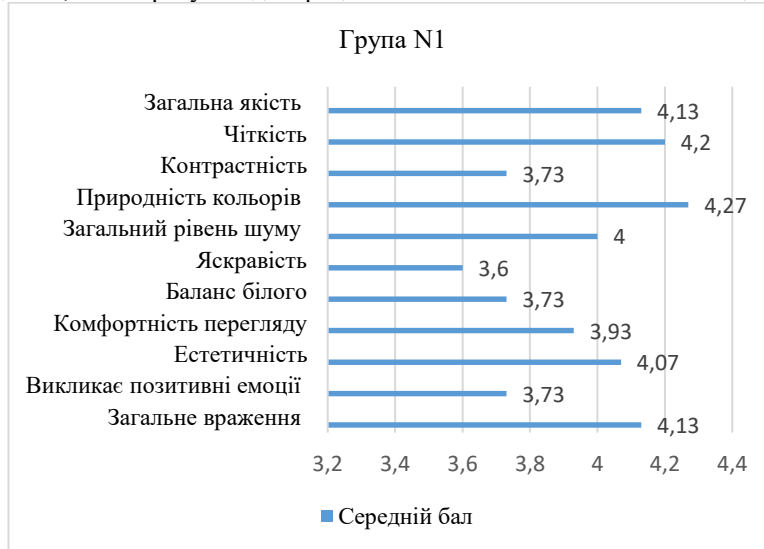


Рис.2. Оцінювання якості зображення групою N1

Група N2 (рис. 3), до якої входили професіонали в сфері зображення, продемонструвала доволі збалансовані показники. Найвищу оцінку отримав показник природності кольорів (4.4), який впливає на загальне сприйняття зображення. Підтвердженням цього є високі оцінки загальної якості зображення (4.13), загального враження від зображення (4) та комфортності перегляду (4.07).



Рис.4. Оцінювання якості зображення групою N2

Це свідчить про важливість відповідності технічним і професійним вимогам до зображень. Однакову кількість балів набрали показники яскравості, балансу білого та наявності шуму на зображенні (3.67). Низькі оцінки цих показників можуть свідчити про менш значущий вплив цих критеріїв для цієї групи респондентів або про недоліки в зображеннях. Інші показники, такі як естетичність, здатність викликати позитивні емоції отримали хоч і нижчі показники, але вони вищі за середню оцінку. Це підкреслює емоційне сприйняття та чутливість професійних фотографів (відеографів) до цих аспектів.. Найнижче оцінено контрастність (3.4). Приблизно збалансованість показників групи N2 відображає їхній професійний підхід і цілісне сприйняття зображення.

На основі представленого графіка для групи N3 (рис. 4) можна зробити такі висновки: природність кольорів (4.6) і баланс білого (4.73) є одними з найбільш високо оцінених характеристик, що може свідчити про важливість візуальної достовірності та природності кольорів.

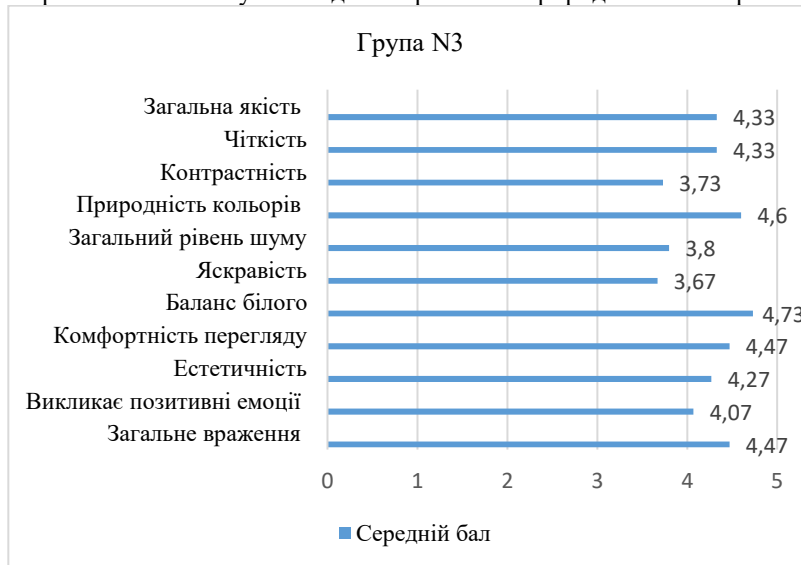


Рис.4. Оцінювання якості зображення групою N3

Показник комфортності перегляду (4.47) також набрав високої оцінки, що підкреслює значення зручності та приємності взаємодії із зображенням. Високі оцінки отримали загальне враження (4.47) і загальна якість (4.33), що демонструє позитивне сприйняття зображення серед респондентів цієї групи. Дещо нижчу оцінку отримали естетичність (4.27) та здатність зображення викликати позитивні емоції (4.07), що свідчить про загальну задоволеність виглядом зображень. Нижчі оцінки мають контрастність (3.73) і загальний рівень шуму (3.8) порівняно з іншими характеристиками, що може вказувати на менш критичне сприйняття цих параметрів цією групою. Загалом респонденти групи N3 демонструють збалансоване сприйняття якості зображень, з акцентом на природність кольорів, баланс білого, комфорт перегляду та загальні враження. Технічні аспекти, такі як чіткість і деталізація, також важливі, але менш критичні, ніж емоційна складова та зручність.

Група N4 (рис. 5) найвищими оцінками позначила природність кольорів (4.53), комфортність перегляду (4.33), загальне враження (4.13), загальну якість (4.07); контрастність, загальний рівень шуму і баланс білого отримали однаковий середній показник (4), інші показники отримали оцінку трохи нижчу, але вона наближена до високих показників. Найнижче респонденти оцінили здатність викликати позитивні емоції (3.53). Група N4 продемонструвала менш критичний підхід до оцінювання, ніж професіонали, однак високо оцінювала характеристики, які безпосередньо впливають на комфорт перегляду.

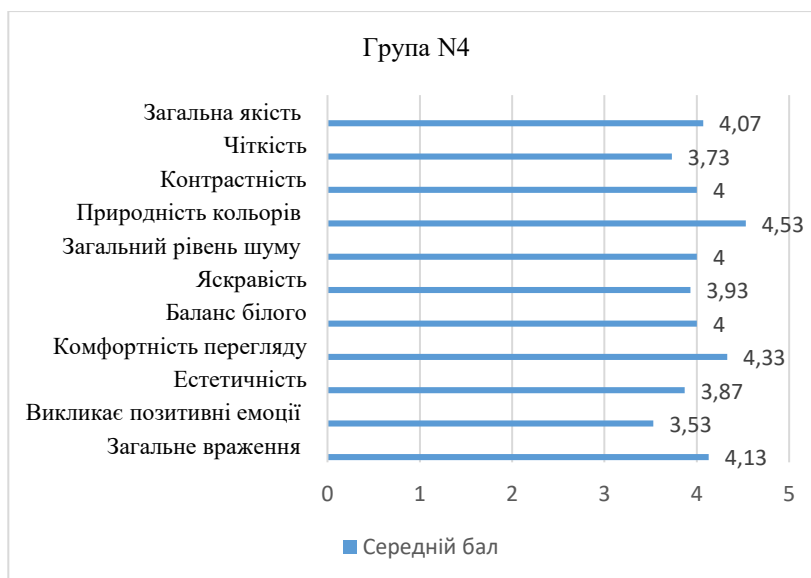


Рис.5. Оцінювання якості зображення групою N4

Узагальнені результати за всіма групами демонструють таке (рис. 6):

Порівняльне оцінювання груп

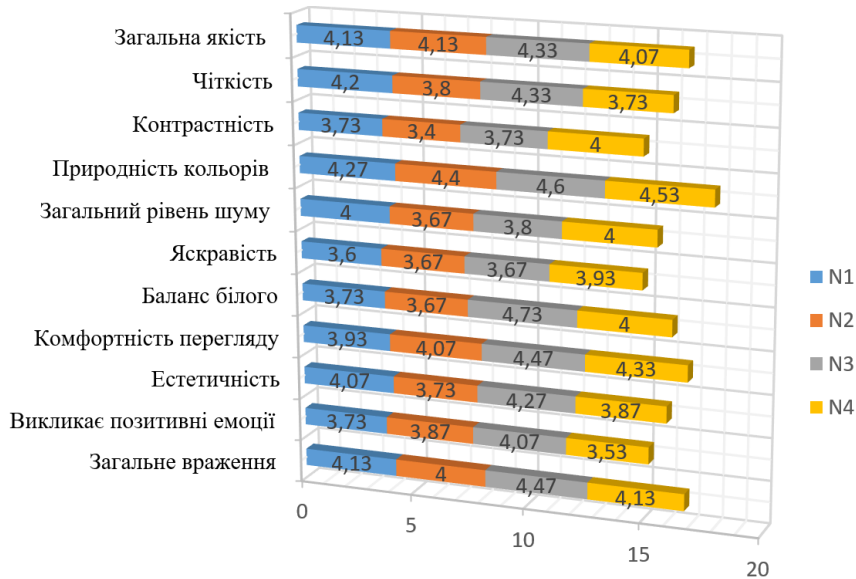


Рис.6. Порівняльне оцінювання груп

1. Оцінка загальної якості зображень. Середній бал від 4.07 до 4.33 вказує на те, що більшість учасників оцінили зображення як «хороше» або «дуже хороше». Причому група N2 та група N3 надавали трохи вищі оцінки, ніж групи N1 та група N4.

2. Оцінка чіткості зображення. Середній бал від 3.73 до 4.33. Найвищий середній бал був у групі N3, що свідчить про те, що ця група краще сприймала рівень деталізації. Учасники більш чутливо оцінювали дрібні деталі, що підтверджує значущість чіткості для суб'єктивного сприйняття. Важливо для користувачів, які цінують деталізацію в центральних і периферійних частинах зображення.

3. Оцінка контрастності зображення. Середній бал від 3.4 до 4.0. Незважаючи на те, що для дослідження вибрано еталонне зображення із яскравими та добре збалансованими кольорами, учасники вказували на певні проблеми із занадто низькою контрастністю.

4. Оцінка природності кольорів. Середній бал від 4.27 до 4.6. Природність кольорів є одним із ключових факторів, що впливає на естетичну оцінку. Усі групи дали максимально високі оцінки природності кольорів. Це означає, що зображення демонструє збалансовану насиченість кольорів, відсутність перенасичених або блідих відтінків, які часто знижують оцінки, тобто кольори виглядають природно, є близькими до реальних.

5. Оцінка загального рівня шуму (зернистості) на зображенні. Середній бал від 3.67 до 4.0. Еталонність зображення не завадила зауважити деякі проблеми з шумом, але здебільшого рівень шуму був прийнятним.

6. Оцінка яскравості. Середній бал від 3.6 до 3.93. Усі групи респондентів вважають зображення доволі яскравим.

7. Оцінка балансу білого. Середній бал від 3.67 до 4.73. Зображення з коректним балансом білого, тому не випадково воно отримало високі оцінки, особливо у групі N2.

8. Комфортність перегляду. Середній бал від 3.93 до 4.47. Надаючи високі оцінки цьому показнику, респонденти вказують на те, що зображення є комфортним для перегляду. Такі зображення можуть використовувати для друку, реклами, соцмереж тощо.

9. Оцінка естетичності зображення. Середній бал від 3.73 до 4.27. Групи N1 і N3 відзначили, що зображення є дуже естетичним.

10. Оцінка здатності зображення викликати позитивні емоції. Середній бал від 3.53 до 4.07.

11. Оцінка загального враження від зображення. Середній бал від 4 до 4.47. Усі групи високо оцінили зображення в цілому. Рівень якості, безперечно, залежить від балансу білого, природності кольорів і відсутності артефактів.

Загалом природність кольорів і баланс білого є ключовими критеріями, які впливають на загальне враження. Чіткість і деталізація важливі для професіоналів, але менш критичні для звичайних користувачів. Респонденти помітили, що зображення з надмірним шумом і слабким контрастом сприймаються менш якісними. Професіонали більше звертають увагу на технічні деталі, звичайні користувачі орієнтуються на загальне враження та естетику. Позитивні емоції від зображення теж мають високу оцінку у всіх групах, що свідчить про важливість загального візуального ефекту.

Найнижчі оцінки стосуються балансу білого та контрастності, що свідчить про загальні проблеми у сприйнятті цих характеристик серед усіх респондентів.

Група N2 виявилась найвимогливішою, що відповідає професійному рівню респондентів, а група N4 – найменш критичною. Що теж підтвердив факт, оскільки респонденти не володіють технікою професійної корекції зображень.

Серед переваг суб'єктивного підходу з одним стимулом можемо відзначити такі:

- оцінка естетичних та емоційних аспектів: суб'єктивні методи дозволяють оцінити, крім технічних параметрів, художню привабливість зображень для респондентів чи природність кольорів, які складно оцінити за допомогою математичних моделей;

- виявлення прихованих дефектів: респонденти можуть помітити дефекти, які не враховано об'єктивними метриками, наприклад, артефакти стиснення, незначні спотворення кольорів або незвичні текстури. Суб'єктивні методи дають змогу ідентифікувати ці недоліки, що може бути критично важливим для високоякісної фотографії та відео;

- підвищення реалістичності зображень: моделі, навчені на основі суб'єктивних оцінок, здатні створювати зображення, які виглядають більш природно і привабливо для користувачів;

- диференціація за групами: респонденти з різним професійним досвідом по-різному оцінюють якість зображення;

- універсальність оцінювання: усі групи відзначили гармонію кольорів і реалістичність як універсальні критерії, які можуть позитивно впливати на якість сприйняття.

Серед недоліків суб'єктивного підходу з одним стимулом можемо відзначити такі:

- часозатратність: організація суб'єктивного тестування, збір даних потребують значного часу для залучення респондентів, підготування матеріалів, проведення опитування та його аналізу;

- ресурсозатратність: для забезпечення надійності експерименту потрібно залучити значну кількість респондентів;

- індивідуальність сприйняття: на суб'єктивність оцінки зображень можуть впливати індивідуальні особливості респондентів (вік, досвід, освіченість, професія, культурний рівень тощо). Наприклад, те, що для одного глядача виглядає гармонійно, може здаватися іншому неприродним.

- психологічні чинники: на оцінку якості зображення може впливати емоційний стан респондентів на момент опитування, зокрема ефект втоми, ефект першого враження; складність відокремлення факторів (наприклад, респондент може оцінити низьку якість зображення через невдалу композицію, а не технічні характеристики);

- технічне забезпечення: оскільки для респондентів було обрано реальні умови у форматі онлайн, то в такому разі всіх учасників тестування неможливо забезпечити однаковими умовами. Насамперед, через відсутність уніфікованого обладнання для тестування. Респонденти проходили тестування з різних пристроїв, відповідно були екрани різної якості. Незважаючи на приблизно однаковий час проходження тестування, освітлення чи контрастність середовища могли спотворювати сприйняття;

- проблеми з масштабуванням: для мінімального набору, на відміну від великих наборів даних (тисячі зображень), суб'єктивний підхід є доволі практичним;

- проблема з повторюваністю: результати суб'єктивного тестування важко відтворити, оскільки вони залежать від конкретної групи респондентів.

Огляд основних концепцій і методів демонструє необхідність інтеграції суб'єктивних оцінок із технологіями глибокого навчання для покращення оцінки якості зображень. Це сприятиме створенню адаптивних моделей, які краще відповідають людському сприйняттю.

Висновки

Суб'єктивний метод оцінки якості зображення з одним стимулом – це підхід, який відображає комплексне сприйняття візуального контенту, враховуючи і технічні, й естетичні аспекти. На відміну від об'єктивних метрик, суб'єктивна оцінка дозволяє враховувати різні психологічні і культурні фактори, які впливають на сприйняття зображення. Оцінювання якості зображення безпосередньо глядачами дозволяє отримати найбільш релевантні дані щодо того, як зображення сприймають у реальних умовах використання. Створення універсальної моделі може ускладнити наявність у різних груп користувачів різних критеріїв щодо оцінювання якості зображень.

Суб'єктивне тестування за допомогою методу з одним стимулом виявилось ефективним для виявлення сильних і слабких сторін зображення, а також для розуміння, які аспекти є пріоритетними для різних груп респондентів. Розгляд основних концепцій та методів підкреслює важливість інтеграції суб'єктивних оцінок із сучасними технологіями глибокого навчання для створення моделей, що більш точно відображають реальне сприйняття якості зображень.

Моделі, які навчаються на основі суб'єктивних даних, можуть бути більш точними в оцінці таких аспектів, як реалістичність, естетична привабливість, видимість артефактів тощо. Це відкриває нові горизонти для досліджень у сфері оброблення зображень, особливо у контексті персоналізації контенту. Інтеграція суб'єктивних оцінок із глибоким навчанням дозволяє вирішити проблему невідповідності між об'єктивними метриками та людським сприйняттям, створити комбіновані метрики, які зможуть більш точно відобразити якість зображень.

Література

1. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13 (4), 600-612. doi: 10.1109/TIP.2003.819861
2. Chandler, D.M. (2013) Seven Challenges in Image Quality Assessment: Past, Present, and Future Research. *ISRN Signal Processing*, 1-53. doi: [10.1155/2013/905685](https://doi.org/10.1155/2013/905685)
3. Романишин, Ю., Єлманов, С., Андрухів, Т., Теглівець О. (2023) Оцінка якості зображень згортковою нейронною мережею при використанні бази tid2013. *ICTEE*, випуск 3, номер 2, 170-179. doi: [10.23939/ictee2023.02.170](https://doi.org/10.23939/ictee2023.02.170)
4. Ledig C. et al. (2017) Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *CVPR*, 105-114. doi: 10.1109/CVPR.2017.19
5. Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L. (2017) Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, 26, 3142-3155. doi: [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206)
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial networks. *Advances in neural information processing systems*. doi: 10.1145/3422622
7. Dong, C., Loy, C.C., He, K. and Tang, X. (2016) Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38, 295-307. doi:10.1109/TPAMI.2015.2439281
8. Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. doi: 10.48550/arXiv.1511.06434
9. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric, IEEE/CVF Conference on Computer Vision and Pattern Recognition, 586-595. doi:[10.1109/CVPR.2018.00068](https://doi.org/10.1109/CVPR.2018.00068)
10. Yanagawa, M., Hata, A., Honda, O. et al. (2018) Subjective and objective comparisons of image quality between ultra-high-resolution CT and conventional area detector CT in phantoms and cadaveric human lungs. *Eur Radiol*, 28, 5060–5068. doi:[10.1007/s00330-018-5491-2](https://doi.org/10.1007/s00330-018-5491-2)
11. Talebi, H., & Milanfar, P. (2018). NIMA: Neural image assessment. *IEEE Transactions on Image Processing*. vol. 27, no. 8, 3998-4011, doi: 10.1109/TIP.2018.2831899.
12. Bovik, A.C., Adams, J.E. (2009) The Essential Guide to Image Processing. *J. Electronic Imaging*, 19, 029901. doi:10.1117/1.3430078
13. Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8), 2378-2386. doi:10.1109/tip.2011.2109730
14. Hu, B., Li, L., Wu, J., Qian, J. (2020) Subjective and objective quality assessment for image restoration: A critical survey/ *Signal Processing: Image Communication*, 85, 115839. doi:[10.1016/j.image.2020.115839](https://doi.org/10.1016/j.image.2020.115839)
15. Khodabakhsh, A., Pedersen, M., Busch, C. (2019) Subjective Versus Objective Face Image Quality Evaluation For Face Recognition ICBEA, *Proceedings of the 2019 3rd International Conference on Biometric Engineering and Applications*, pp. 36-42. doi: [10.1145/3345336.3345338](https://doi.org/10.1145/3345336.3345338)
16. Winkler, S. (2013). Digital video quality: Vision models and metrics. 200 p.
17. Zhang, Fan & Li, Songnan & Ma, Lin & Ngan, King. (2010). Limitation and challenges of image quality measurement. *Proceedings of SPIE - The International Society for Optical Engineering*. doi: 7744. 10.1117/12.863083.
18. ITU-T REC P.910 (2008): Subjective audiovisual quality assessment methods for multimedia applications.
19. Chollet F. (2021) Deep Learning with Python, Manning; 2nd ed. 504 pp.
20. Talun, A., Drozda, P., Yelmanov, S., Romanyshyn, Y., Tehlivets, O. (2023) Convolutional Neural Network Assessment of Image Quality Based on the TID2013 Database. *IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS*, 1088–1092. doi: 10.1109/IDAACS58523.2023.10348876
21. Mantiuk, R.K.; Tomaszewska, A.; Mantiuk, R (2012) Comparison of four subjective methods for image quality assessment. *Computer Graphics Forum*, 31(8), pp. 2478-2491. Doi: 10.1111/j.1467-8659.2012.03188.x
22. Marni, F., Bertini, M., Galteri, L. and Del Bimbo, A., (2021) A NoGAN approach for image and video restoration and compression artifact removal, *25th International Conference on Pattern Recognition (ICPR)*, pp. 9326-9332, doi: 10.1109/ICPR48806.2021.9413095
23. Yu, X. et al., (2022) Subjective Quality Assessment of User-Generated Content Gaming Videos, *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, Waikoloa, HI, USA, 74-83, doi: 10.1109/WACVW54805.2022.00013.
24. Yuan P, Bai R, Yan Y, Li S, Wang J, Cao C, Wu Q. (2023) Subjective and objective quality assessment of gastrointestinal endoscopy images: From manual operation to artificial intelligence. *Front Neurosci. Feb 14;16:1118087*. doi: 10.3389/fnins.2022.1118087.
25. Singh, O.P., Singh, K.N., Singh, A.K., Agrawal, A.K. (2025) NeurocomputingDeep learning-based image encryption techniques: *Fundamentals, current trends, challenges and future directions*, 612, 128714. doi: 10.1016/j.neucom.2024.128714

26. Talun, A., Drozda, P., Romanyshyn, Y., Tehlivets, O., Yelmanov, S. (2023) Test Images for Training Convolutional Neural Networks for Image Contrast Assessment. *International Scientific and Technical Conference on Computer Sciences and Information Technologies*. doi: 10.1109/CSIT61576.2023.10324137
27. Virtanen T., Nuutinen M., Vaahteranoksa M., Oittinen P. and J. Häkkinen (2015) CID2013: A Database for Evaluating No-Reference Image Quality Assessment Algorithms, *IEEE Transactions on Image Processing*, vol. 24, no. 1, 390-402, doi: 10.1109/TIP.2014.2378061
28. Recommendation ITU-R BT.500-11 Methodology for the subjective assessment of the quality of television pictures.

References

1. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*. 13 (4), 600-612. doi: 10.1109/TIP.2003.819861
2. Chandler, D.M. (2013) Seven Challenges in Image Quality Assessment: Past, Present, and Future Research. *ISRN Signal Processing*, 1-53. doi: [10.1155/2013/905685](https://doi.org/10.1155/2013/905685)
3. Romanyshyn Y., Yelmanov S., Andrukhiv T., Teglivets O. (2023). Image quality assessment by convolutional neural network using the tid2013 database. *ICTEE*, 3(2), 170-179. doi: [10.23939/ictee2023.02.170](https://doi.org/10.23939/ictee2023.02.170)
4. Ledig C. et al. (2017) Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *CVPR*, 105-114. doi: 10.1109/CVPR.2017.19
5. Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L. (2017) Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, 26, 3142-3155. doi: [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206)
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial networks. *Advances in neural information processing systems*. doi: 10.1145/3422622
7. Dong, C., Loy, C.C., He, K. and Tang, X. (2016) Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38, 295-307. doi:10.1109/TPAMI.2015.2439281
8. Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. doi: 10.48550/arXiv.1511.06434
9. Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 586-595. doi: [10.1109/CVPR.2018.00068](https://doi.org/10.1109/CVPR.2018.00068)
10. Yanagawa, M., Hata, A., Honda, O. et al. (2018) Subjective and objective comparisons of image quality between ultra-high-resolution CT and conventional area detector CT in phantoms and cadaveric human lungs. *Eur Radiol*, 28, 5060–5068. doi: [10.1007/s00330-018-5491-2](https://doi.org/10.1007/s00330-018-5491-2)
11. Talebi, H., & Milanfar, P. (2018). NIMA: Neural image assessment. *IEEE Transactions on Image Processing*. vol. 27, no. 8, 3998-4011, doi: 10.1109/TIP.2018.2831899.
12. Bovik, A.C., Adams, J.E. (2009) The Essential Guide to Image Processing. *J. Electronic Imaging*, 19, 029901. doi:10.1117/1.3430078
13. Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8), 2378-2386. doi:10.1109/tip.2011.2109730
14. Hu, B., Li, L., Wu, J., Qian, J. (2020) Subjective and objective quality assessment for image restoration: A critical survey/ *Signal Processing: Image Communication*, 85, 115839. doi: [10.1016/j.image.2020.115839](https://doi.org/10.1016/j.image.2020.115839)
15. Khodabakhsh, A., Pedersen, M., Busch, C. (2019) Subjective Versus Objective Face Image Quality Evaluation For Face Recognition ICBEA, *Proceedings of the 2019 3rd International Conference on Biometric Engineering and Applications*, pp. 36-42. doi: [10.1145/3345336.3345338](https://doi.org/10.1145/3345336.3345338)
16. Winkler, S. (2013). Digital video quality: Vision models and metrics. 200 p.
17. Zhang, Fan & Li, Songnan & Ma, Lin & Ngan, King. (2010). Limitation and challenges of image quality measurement. *Proceedings of SPIE - The International Society for Optical Engineering*. doi: 7744. 10.1117/12.863083.
18. ITU-T REC P.910 (2008): Subjective audiovisual quality assessment methods for multimedia applications.
19. Chollet F. (2021) *Deep Learning with Python*, Manning; 2nd ed. 504 pp.
20. Talun, A., Drozda, P., Yelmanov, S., Romanyshyn, Y., Tehlivets, O. (2023) Convolutional Neural Network Assessment of Image Quality Based on the TID2013 Database. *IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS*, 1088–1092. doi: 10.1109/IDAACS58523.2023.10348876
21. Mantiuk, R.K.; Tomaszewska, A.; Mantiuk, R (2012) Comparison of four subjective methods for image quality assessment. *Computer Graphics Forum*, 31(8), pp. 2478-2491. Doi: 10.1111/j.1467-8659.2012.03188.x
22. Mameli, F., Bertini, M., Galteri, L. and Del Bimbo, A., (2021) A NoGAN approach for image and video restoration and compression artifact removal, *25th International Conference on Pattern Recognition (ICPR)*, pp. 9326-9332, doi: 10.1109/ICPR48806.2021.9413095
23. Yu, X. et al., (2022) Subjective Quality Assessment of User-Generated Content Gaming Videos, *IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, Waikoloa, HI, USA, 74-83, doi: 10.1109/WACVW54805.2022.00013.
24. Yuan P, Bai R, Yan Y, Li S, Wang J, Cao C, Wu Q. (2023) Subjective and objective quality assessment of gastrointestinal endoscopy images: From manual operation to artificial intelligence. *Front Neurosci*. Feb 14;16:1118087. doi: 10.3389/fnins.2022.1118087.
25. Singh, O.P., Singh, K.N., Singh, A.K., Agrawal, A.K. (2025) Neurocomputing Deep learning-based image encryption techniques: *Fundamentals, current trends, challenges and future directions*, 612, 128714. doi: 10.1016/j.neucom.2024.128714
26. Talun, A., Drozda, P., Romanyshyn, Y., Tehlivets, O., Yelmanov, S. (2023) Test Images for Training Convolutional Neural Networks for Image Contrast Assessment. *International Scientific and Technical Conference on Computer Sciences and Information Technologies*. doi: 10.1109/CSIT61576.2023.10324137
27. Virtanen T., Nuutinen M., Vaahteranoksa M., Oittinen P. and J. Häkkinen (2015) CID2013: A Database for Evaluating No-Reference Image Quality Assessment Algorithms, *IEEE Transactions on Image Processing*, vol. 24, no. 1, 390-402, doi: 10.1109/TIP.2014.2378061
28. Recommendation ITU-R BT.500-11 Methodology for the subjective assessment of the quality of television pictures.