



**KAPITEL 4 / CHAPTER 4<sup>4</sup>**  
**METHODS OF INTELLECTUAL IMAGE AND VIDEO PROCESSING  
BASED ON VISUAL QUALITY METRICS FOR EMERGING  
APPLICATIONS**

**DOI: 10.30890/2709-2313.2023-20-02-010**

## **Вступ**

Зображення та відео, що отримуються різноманітними типами датчиків та систем, широко використовуються в даний час для вирішення різних важливих проблем у багатьох додатках: у промисловості, картографуванні земельних покривів, сільському та лісовому господарстві, контролі якості, робототехніці, екологічному моніторингу тощо [1-3]. Однією із сучасних тенденцій є використання багатоканальних систем, а саме кольорових, мульти- та гіперспектральних, багатопоглядових (наприклад, стерео) тощо [4, 5]. Відео також можна розглядати як багатоканальні (багатокадрові) дані. Ще одна тенденція полягає в оцінці та врахуванні візуальної якості зображень, що приймається до уваги на етапах отримання та обробки зображень. Під обробкою будемо надалі розуміти широкий набір операцій, включаючи оцінку характеристик шуму або викривлень, фільтрацію (придушення шуму), стиснення, інтерпретацію та оцінку параметрів об'єкта, прийняття рішення.

Це потребує розв'язання широкого набору практичних завдань. Однією з них є розробка та тестування візуальних показників (метрик) якості, які базуються на врахуванні особливостей системи зору людини (СЗЛ) [6-8]. Незважаючи на значні зусилля багатьох дослідників, витрачені на розробку показників якості на основі СЗЛ, навіть найкращі з них все ще страждають від істотних недоліків. Зокрема, деякі з них недостатньо універсальні. Крім того, часто незрозуміло, які показники на основі СЗЛ є найкращими для конкретних застосувань. Однією з причин цього є те, що валідація метрик є трудомістким завданням і потрібні належним чином розроблені бази даних спотворених зображень або відеопослідовностей. Комбіновані метрики є прийнятним рішенням, але їх розробка ще далека від завершення. Вже створені бази даних можуть бути використані в деяких випадках, але створення нових баз даних із конкретними типами та рівнями спотворень також є бажаним для деяких

---

<sup>4</sup>*Authors: Lukin V.V., Ieremeiev O.I., Abramova V.V., Abramov S.K., Rubel O.S., Krzysztof O., Piotr L., Yaroslav F.*



додатків, таких, як контроль якості в промисловості, аналіз текстурованих матеріалів, дистанційне зондування, зображення з автоматизованих рухомих транспортних засобів як автономні автомобілі, безпілотні літальні апарати (БПЛА) та дрони.

Ще одне завдання - це включення метрик якості (МЯ) на основі СЗЛ в обробку зображень/відео. Зокрема, бажано в повній мірі використовувати МЯ на основі СЗЛ в розробці фільтрів, аналізі характеристик розглянутих фільтрів, підборі фільтра та налаштуванні його параметрів, прийнятті рішення щодо доцільності використання етапу придушення шуму [9]. Це ж стосується стиснення зображень із втратами [10-15]. Зазвичай пікове співвідношення сигнал-шум (ПССШ) використовується як єдиний параметр, що характеризує ефективність кодера для різних коефіцієнтів стиснення. Тим часом, якість візуального зображення не відображена належним чином в ПССШ. Отже, слід використовувати інші показники. Їх значення слід передбачити перед стисненням та/або забезпечити на етапі стиснення [14-18]. Крім того, бажаною є розробка високошвидкісних автоматизованих методів обробки зображень та відео. Звичайно, можна залучити експертів на деяких етапах обробки, але такі можливості обмежені.

Метою спільного проекту була розробка методів та алгоритмів інтелектуальної обробки зображень та відео на основі метрик візуальної якості для нових застосувань. В цьому розділі головну увагу приділено розробці та верифікації так званих комбінованих показників (метрик візуальної якості), які використовують позитивні риси декількох "різномірних" елементарних МЯ на основі СЗЛ. В такому випадку швидка обробка зображення досягається за рахунок паралелізації обчислення елементарних МЯ та досить простого «агрегування» елементарних результатів [19-21]. Агрегування виконувалось одним з трьох методів: 1) за допомогою навчених нейронних мереж [19]; 2) з використанням попередньої лінеаризації та стійкої обробки значень метрик [20]; 3) на основі перемноження метрик з оптимізацією вагових коефіцієнтів [21].

У перших двох випадках розглядається задача розробки комбінованих метрик для їх застосувань у обробці даних багатоканального (зокрема триканального) дистанційного зондування. У останньому випадку йдеться про промислове застосування – виготовлення об'єктів за допомогою 3D принтера.



#### 4.1. Метод стійкого комбінування метрик для систем ДЗ

Найпоширенішими системами дистанційного зондування (ДЗ) є такі: оптичні та радіолокаційні сенсори; мультиспектральні та гіперспектральні датчики та інші [22]. Кожен тип системи ДЗ має певні переваги та недоліки. Радари, наприклад, дозволяють отримувати унікальні дані про підповерхневі об'єкти і можуть здійснювати моніторинг у поганих погодних умовах [23].

Ці системи використовують різні принципи формування (отримання) зображення, але всі вони, по суті, об'єднані проблемою наявності шуму та спотворень. Подальше застосування певних методів цифрової обробки, таких як фільтрація [24, 25] або стиснення з втратами [12, 26, 27], може як зменшити рівень спотворення даних, так і призвести до більшої деградації та втрати важливої інформації. Тому доцільно застосовувати оцінку якості зображення, щоб визначити ступінь погіршення зображення для вихідних даних та, як результат обробки зображення, запобігти небажаному рівню втрати корисної інформації.

Враховуючи великий об'єм даних ДЗ та автоматизацію обробки, доречним є використання об'єктивних методів оцінки якості зображення (image quality assessment - IQA) для контролю якості первинних зображень [6] та змін якості, що вносяться в результаті обробки зображень [28]. Є питання щодо того, які методи IQA необхідно використовувати при дистанційному зондуванні [6, 28-30]. У цьому розділі ми розглядаємо критерії (метрики), які стосуються візуальної якості зображень ДЗ. Це важливо, як мінімум, з двох причин. По-перше, зображення ДЗ часто візуалізуються та аналізуються людьми. Тоді, враховуючи особливості системи зору людини (human vision system - HVS), важливо забезпечити збереження інформації на всіх етапах обробки. По-друге, показники візуальної якості стосуються ефективності таких операцій, як виявлення малорозмірних об'єктів, класифікація текстурних ділянок тощо [31]. Це означає, що варто використовувати візуальні показники якості для аналізу вихідної якості зображення ДЗ та ефективності обробки даних ДЗ (наприклад, шумопридушення, стиснення з втратами, реконструкції).

Тим часом використання візуальних показників якості у зображеннях ДЗ все ще обмежене. Деякі приклади - це нові HVS-метрики для посилення різкості зображення [32] та їх суміщення [33, 34], а також розпізнавання об'єктів [35]. Однак стара HVS-метрика SSIM все ще є популярною для опису якості



зображення [36, 37], хоча відомо, що SSIM не є однією з найкращих метрик візуальної якості.

Протягом останніх десятиліть було розроблено та запропоновано велику кількість метрик візуальної якості з еталоном [38, 39]. Методи IQA на їх основі суттєво відрізняються своєю точністю (адекватністю). Тому загальноприйнятим підходом є перевірка (верифікація) метрик на спеціальних базах даних тестових зображень, таких як TID2013 [40], KADID10k [41], MDID [42] та інших. У таких базах даних для кожного зображення суб'єктивна оцінка візуальної якості (осереднена думка чи оцінка, mean opinion score - MOS) отримується в результаті експериментів з багатьма спостерігачами. Точність чи адекватність метрик оцінюється шляхом порівняння значень метрик та MOS за допомогою критеріїв кореляції. Для цієї мети широко використовуються коефіцієнт кореляції Пірсона (Pearson correlation coefficient - PCC) та ранговий коефіцієнт кореляції Спірмена (Spearman rank order correlation coefficient - SROCC). Зверніть увагу, що для перелічених баз даних зображень точність більшості показників згідно із SROCC не перевищує 0,9.

Щодо зображень ДЗ, то на даний момент не існує спеціально підготовлених баз даних із оцінками MOS. Однак, на наш погляд, існуючі бази даних кольорових зображень можуть бути використані для початкового аналізу та проектування. Тут слід мати на увазі, що ці бази даних містять типи спотворень, що спостерігаються і на зображеннях ДЗ. Тим часом існують також типи спотворень, які не зустрічаються в зображеннях ДЗ. Таким чином, дані з існуючих баз даних слід вибирати ретельно.

Як уже зазначалося, існуючі метрики мають свої обмеження та недоліки. Спільне використання декількох метрик дозволяє компенсувати недоліки одних метрик за рахунок інших. Таким чином, різні комбінаційні підходи можуть підвищити точність оцінки візуальної якості. Деякі автори, зокрема польські співавтори цього розділу запропонували застосовувати комбінацію перемножених функцій із ваговими коефіцієнтами [43]. Інший підхід, запропонований українськими вченими раніше в [44], передбачає використання стійких методів для отримання кінцевого значення комбінованої метрики. Для вирішення цієї проблеми можуть також використовуватися штучні нейронні мережі, які широко використовуються в різних областях [45].

Створення комбінованої метрики передбачає визначення деяких оптимальних значень параметрів та комбінацію метрик, щоб забезпечити



максимально можливий коефіцієнт кореляції для даної бази зображень. Через значні відмінності у змісті та методології суб'єктивної оцінки бази даних зображень можуть суттєво відрізнятись одна від одної. Тому занадто точна оптимізація для конкретної бази даних зображень може призвести до ефекту "перенавчання", що знизить ефективність роботи з іншими наборами зображень.

Мета роботи, результати якої викладені у даному розділі, є потрійною. Спочатку ми хотіли б проаналізувати поведінку метрик для набору спотворень, типових для дистанційного зондування. По-друге, ми аналізуємо використання методу стійкого комбінування [44], який зарекомендував себе як дуже стабільний, для створення вдосконаленої комбінованої метрики якості зображення, оптимізованої для систем дистанційного зондування. По-третє, бажано перевірити, наскільки ефективна ця метрика для різних баз даних.

Загалом, спотворення на зображеннях дистанційного зондування можуть бути різного типу та інтенсивності. Крім того, вони можуть мати різну інтенсивність у компонентних зображеннях багатоканальних даних ДЗ [22, 24, 25]. Виникає питання, як агрегувати результати IQA для багатоканальних зображень ДЗ, якщо IQA було зроблено окремо для компонентних зображень або їх груп. Давайте відповімо на це питання пізніше і спочатку розглянемо простіші випадки спотворених компонентних зображень або трикомпонентних зображень із викривленнями.

Тут треба коротко розглянути типові спотворення, які можуть виникати під час отримання та обробки зображень при дистанційному зондуванні. Основними їх видами є наступні:

– адитивний білий гаусів шум, який можна розглядати як (ідеалізовану) модель шуму для багатьох типів зображень ДЗ, таких як, наприклад, мультиспектральні [22, 24]; зауважимо, що шум може бути різної інтенсивності в різних компонентах багатоканальних зображень [24];

– просторово корельований шум, який може бути присутнім у вихідних даних або виникати в результаті інтерполяції даних на етапі реєстрації зображення [22, 25];

– різні замасковані або високочастотні шуми, які можуть бути присутніми на текстурних зображеннях або з'являтися після усунення розмитості [22];

– залежний від сигналу (зокрема, мультиплікативний) шум, характерний для гіперспектральних та радіолокаційних зображень [28];



- імпульсний шум, який може з'явитися після передачі зображення по лініях зв'язку або внаслідок інших факторів [22];
- шум квантування, який може виникнути на етапі оцифрування даних або з'явитися внаслідок деяких операцій нормалізації або перетворення [22];
- розмиття зображення, яке може бути результатом руху системи дистанційного зондування або неоднорідності середовища поширення сигналу [22];
- спотворення внаслідок стиснення з втратами (наприклад, JPEG або JPEG2000), що може суттєво зменшити розмір файлу, але вносить погіршення, які можуть бути помітними або навіть дратувати [26, 48];
- спотворення, спричинені методом, який усуває шум, але одночасно та неминуче призводить до погіршення інформаційної складової [28];
- складні спотворення, наприклад, у разі стиснення з втратами зображень, що спотворені шумом [26].

Зазначимо, що всі ці типи спотворень можна моделювати, а зображення дистанційного зондування різного походження можуть (на початковому етапі досліджень) бути замінені на більш поширені оптичні зображення з відповідними спотвореннями. Тут ми дійшли до необхідності проаналізувати існуючі бази даних спотворених компонентних та кольорових зображень. Зауважимо, що в багатьох базах даних розглядається обмежена кількість типів шуму та спотворень. Більшість із них зосереджені на врахуванні таких типових простих спотворень, як адитивні шуми, розмитість, спотворення внаслідок стиснення JPEG та JPEG2000. Для розглянутого завдання потрібні більш складні та великі бази даних зображень. Найбільш підходящі з них представлені в таблиці 1.

База зображень TID2013 містить 25 еталонних та 3000 спотворених зображень з 24 типами спотворень. Опублікована у 2019 році база даних KADID10k містить понад 10000 зображень із 25 типами спотворень. Інша база - MDID [24] - складається з 1600 зображень із комбінованими (множинними) спотвореннями. Числа у стовпцях таблиці 1 для розглянутих баз даних представляють показники спотворень. Знак «-» показує відсутність, а знак «+» вказує на наявність даного типу спотворень у базі даних MDID.



Таблиця 1 – Аналіз типів зображень у базах

##	Тип спотворень	Набір спотворень	Індекс у базі TID2013	Індекс у базі KADID10k	Наявність у MDID
1	Адитивний гаусів шум	"Noise", "Actual"	1	11	+
2	Адитивний шум головним чином в кольорових компонентах	"Noise"	2	12	-
3	Просторово-корельований шум	"Noise", "Actual"	3	-	-
4	Маскований шум	"Noise", "Actual"	4	-	-
5	Високочастотний шум	"Noise", "Actual"	5	-	-
6	Імпульсний шум	"Noise", "Actual"	6	13	-
7	Шум квантування	"Noise"	7	6	-
8	Гаусове розмиття	"Noise", "Actual"	8	-	+
9	Фільтроване зображення	"Noise", "Actual"	9	15	-
10	JPEG стиснення	"Actual"	10	10	+
11	JPEG2000 стиснення	"Actual"	11	9	+
12	Мультиплікативний гаусів шум	"Noise", "Actual"	19	14	-
13	Стиснення з втратами зображень з шумом	"Noise", "Actual"	21	-	-

Як бачимо, база TID2013 надає найширші можливості для проектування та перевірки метрик для цілей дистанційного зондування. Тому ми обрали її для своїх досліджень. Викривлення в TID2013 були згруповані в підмножини [40] відповідно до їх типів. Між ними підмножини "Noise" та "Actual" можна розглядати як такі, що пов'язані з дистанційним зондуванням. Кольорові та інші спотворення, що є характерними для кольорових зображень, не є типовими для області, що нас цікавить, і тому вони не будуть нами розглядатися.

Протягом останніх років було розроблено велику кількість нових метрик. Нами вивчено п'ятдесят з них, а результати для обраних підмножин спотворень наведені в таблиці 2. Детальну інформацію про показники в таблиці 2 можна отримати за таким посиланням [46]. Ефективність метрик характеризується SROCC, оскільки цей критерій не залежить від абсолютних значень розглянутих метрик, а лише від їх рангів.



Таблиця 2 – Значення SROCC для елементарних метрик

##	Метрики з еталоном	Всі типи спотворень	Noise	Actual	Noise & Actual	##	Метрики з еталоном	Всі типи спотворень	Noise	Actual	Noise & Actual
1	MDSI	0,8897	0,9275	0,9387	0,9374	26	IWSSIM	0,7775	0,8783	0,8934	0,8937
2	CVSSI	0,8090	0,9248	0,9350	0,9341	27	SSIM4	0,7657	0,8743	0,8887	0,8902
3	MCS D	0,8045	0,9224	0,9326	0,9323	28	CSSIM	0,8417	0,8728	0,8878	0,8892
4	PSNRHA	0,8198	0,9230	0,9388	0,9322	29	MSSIM	0,7872	0,8733	0,8872	0,8881
5	GMSD	0,8004	0,9211	0,9314	0,9318	30	RFSIM	0,7721	0,8731	0,8793	0,8847
6	PSNRHMAm	0,8541	0,9221	0,9387	0,9315	31	VSNR	0,6809	0,8691	0,8817	0,8815
7	PSIM	0,8926	0,9189	0,9309	0,9303	32	ESSIM	0,8121	0,8411	0,8866	0,8619
8	PSNRHAY	0,7794	0,9184	0,9272	0,9275	33	CSSIM4	0,7394	0,8432	0,8668	0,8604
9	PSNRHVS	0,6536	0,9172	0,9257	0,9263	34	GSM	0,8028	0,8408	0,8841	0,8583
10	PSNRHMA	0,8137	0,9151	0,9343	0,9250	35	VIF	0,6816	0,8422	0,8585	0,8532
11	IGM	0,8023	0,9099	0,9220	0,9227	36	NQM	0,6349	0,8362	0,8572	0,8527
12	PSNRHMAy	0,7570	0,9107	0,9209	0,9226	37	RVSIM	0,6748	0,8192	0,8449	0,8423
13	VSI	0,8967	0,9101	0,9258	0,9218	38	MSE	0,6396	0,8217	0,8246	0,8335
14	SRSIM	0,8076	0,9070	0,9211	0,9206	39	PSNR	0,6396	0,8217	0,8246	0,8335
15	HaarPSI	0,8730	0,9063	0,9168	0,9190	40	MSUNIQUE	0,7521	0,7981	0,8276	0,8247
16	ADM	0,7861	0,9113	0,9201	0,9189	41	UNIQUE	0,7466	0,7829	0,8157	0,8117
17	PSNRHVSM	0,6246	0,9061	0,9175	0,9188	42	VIFP	0,6084	0,7835	0,8151	0,8056
18	FSIMc	0,8510	0,9022	0,9150	0,9164	43	CWSSIM	0,5551	0,7943	0,8160	0,8051
19	ADD_GSIM	0,8310	0,9023	0,9151	0,9159	44	DSI	0,7114	0,7493	0,7801	0,7838
20	IQM2	0,7955	0,8995	0,9103	0,9122	45	SSIM	0,6371	0,7574	0,7877	0,7812
21	ADD_SSIM	0,8023	0,9008	0,9119	0,9120	46	IFC	0,5229	0,7201	0,7598	0,7468
22	FSIM	0,8011	0,8969	0,9108	0,9117	47	QILV	0,5975	0,6604	0,6948	0,7049
23	WSNR	0,5796	0,8804	0,8966	0,8952	48	UQI	0,5444	0,6482	0,6904	0,6824
24	SFF	0,8518	0,8787	0,9059	0,8946	49	WASH	0,2903	0,3290	0,4173	0,3139
25	DSS	0,7915	0,8766	0,8904	0,8945	50	MSVD	0,1261	0,1123	0,1309	0,1424

Результати в таблиці 2 представлені в порядку зменшення значення SROCC у стовпці "Noise&Actual". Стовпець "Усі спотворення" стосується всіх 24 типів спотворень у TID2013. Інші стовпці містять дані для підмножин „Noise”, „Actual” та обох.

Як впливає з аналізу даних у колонці "Усі спотворення", немає метрики, яка б забезпечувала SROCC, що перевищує 0,9. Тим часом для підмножин, що нас цікавлять, існує багато метрик якості візуального зображення, які характеризуються  $SROCC > 0,9$ . Найкращі результати мають тенденцію до  $SROCC = 0,94$ . Це оптимістичний результат, який показує, що існує багато універсальних метрик, здатних адекватно характеризувати спотворення на зображеннях ДЗ.

Метрики на верхніх позиціях позначені різними відтінками сірого. Серед них можна вибрати метрики, що забезпечують бажаний компроміс властивостей, зокрема, обчислювально ефективні метрики. У той же час представлені дані можуть служити основою для вибору елементарних метрик (далі ми називаємо елементарними метрики, представлені в табл. 2), які можуть бути використані при проектуванні комбінованих метрик.





Розробка нових метрик є нелегкою справою й вона не гарантує високу ефективність результату. Інше рішення – скомбінувати існуючі метрики таким чином, щоб вони доповнювали одна іншу та давали більш надійні або точні результати. Відповідно до попередніх досліджень щодо синтезу комбінованих метрик [44], можна використовувати медіану трьох (1) або п'яти (2) лінеаризованих метрик або  $\alpha$ -урізане середнє п'яти лінеаризованих метрик (3):

$$M(3) = \text{median}(M_1^{\text{lin}}, M_2^{\text{lin}}, M_3^{\text{lin}}) \quad (1)$$

$$M(5) = \text{median}(M_1^{\text{lin}}, M_2^{\text{lin}}, M_3^{\text{lin}}, M_4^{\text{lin}}, M_{50}^{\text{lin}}) \quad (2)$$

$$At(5) = \frac{1}{3} \sum_{q=2}^4 M_q^{\text{lin}} \quad (3)$$

де  $M_i^{\text{lin}}$  -  $i$ -а елементарна метрика з еталоном після фітінгу та лінеаризації;  
 $M_{(q)}^{\text{lin}}$  -  $q$ -а порядкова статистика.

Для нашого підходу, лінеаризацію метрик до MOS було проведено для бази TID2013. Головна мета цього – не оптимізація для конкретної бази, а можливість представити усі елементарні метрики в єдиному масштабі, бо в первинному вигляді вони мають суттєво різний масштаб (логарифмічний, експоненційний тощо) та діапазони значень.

**Таблиця 3 – Результати для найкращих робасних комбінованих метрик, що були розроблені раніше**

Метод	Елементарні метрики	Всі типи спотворень	Noise	Actual	Noise & Actual
$\alpha$ -урізане середнє	IFC, DCTUNE, SFF, FSIMc, PSNRHMAM	0,8871	0,9219	0,9347	0,9326
Медіана	VIFP, DCTUNE, SFF, FSIMc, PSNRHMAM	0,8863	0,9203	0,9336	0,9314
MDSI		0,8897	0,9275	0,9387	0,9374

Нагадаємо, що підхід [44] дозволив збільшити SROCC приблизно з 0,85 для елементарних метрик до 0,887 для альфа-урізаного середнього значення метрик IFC, DCTUNE, FSIMc, SFF та PSNRHMAM (табл. 3), усі спотворення). Раніше розроблені надійні комбіновані метрики дають хороші результати для наборів Noise та Actual, які лише трохи гірші, ніж для найкращих сучасних метрик з



еталоном (див. результати в нижньому рядку в таблиці 3 для метрики MDSI). Крім того, це означає, що можна очікувати кращих результатів для надійних комбінованих метрик, якщо нові метрики будуть використовуватися як елементарні.

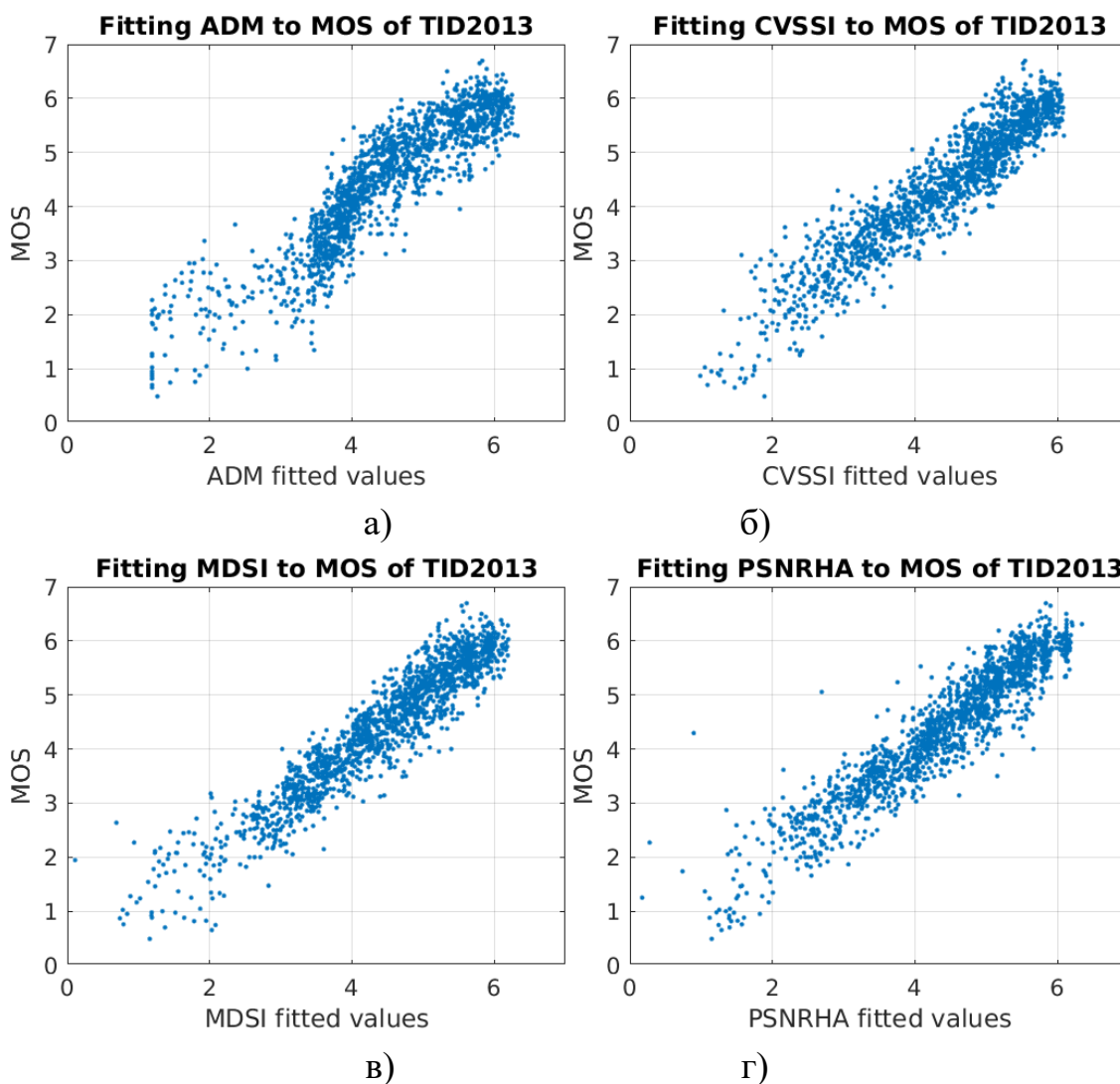
Для застосування таких стійких методів поєднання елементарних метрик необхідна лінеаризація елементарних метрик візуальної якості. Потрібно мати прямий і однозначний зв'язок між метриками та суб'єктивними оцінками якості. Отже, належний перерахунок значень метрик до MOS повинен здійснюватися з використанням монотонних кривих. Аналіз кривих також проводився в [44], підтверджуючи, що функція виду

$$y(x) = a \cdot x^b + c \quad (4)$$

забезпечує достатньо високу точність лінеаризації та, загалом, найменшу середньоквадратичну похибку (root mean square error - RMSE). Фітинг виконувався за допомогою інструменту Curve Fitting Toolbox у програмі MATLAB з увімкненою опцією "Стійка". Скатерограми деяких лінеаризованих показників представлені на рисунку 1. Зверніть увагу, що  $MOS > 6$  у TID2013 відповідає дуже високій візуальній якості спотворених зображень.

Враховуючи велику загальну кількість метрик, параметри лінеаризації наведені в таблиці 4 лише для п'яти елементарних метрик найкращої комбінації (див. деталі нижче).

Ідея стійких комбінованих метрик полягає в мінімізації очевидних помилкових (аномальних) оцінок візуальної якості зображення, що формуються системами дистанційного зондування. Альфа-урізане середнє виключає найменше та найбільше значення оцінок MOS. Медіана кількох оцінок MOS також дозволяє уникнути аномальних оцінок, що надходять з елементарних метрик. Як було показано в [44], серед різних надійних методів комбінування елементарних метрик найкращі результати альфа-урізане середнє. Кількість розглянутих метрик під час комбінування сильно впливає на об'єм обчислень. Тому було накладено обмеження на використання не більше п'яти елементарних метрик. Для будь-яких п'яти з розглянутих 50 метрик вдалося створити 2,2 мільйона унікальних комбінацій.



а) метрика ADM; б) метрика CVSSI;  
 в) метрика MDSI; г) метрика PSNR-НА.

**Рисунок 1 – Скатерограми для кількох метрик після фітінгу**

**Таблиця 4 – Параметри фітінгу для найкращої комбінації**

Метрики	а	б	с	RMSE
ADM	5,208	1,058	1,191	0,5133
CVSSI	-10,9	0,5469	6,385	0,4137
MDSI	-17,94	1,679	6,721	0,3945
PSNRHA	-271,3	-1,147	9,81	0,4057
PSIM	3,922	583,8	1,981	0,4144

Як критерій оптимізації комбінованої метрики зазвичай використовують коефіцієнти кореляції. Такі показники кількісно характеризують зв'язок між значеннями метрик та MOS. Кореляція Пірсона та SROCC є найбільш широко



використовуваними коефіцієнтами кореляції. Істотною перевагою коефіцієнтів рангової кореляції є їх незалежність від значень метрик та MOS, оскільки оцінюються їх ранги. Для лінійного коефіцієнта кореляції (Пірсона) необхідна підгонка. Запропонований метод поєднання метрик також вимагає лінеаризації, тому розраховуються як PCC, так і SROCC.

Результатом нашого дослідження стало значне число комбінацій, і необхідно було вибрати найкращу з них. Одним із ключових питань був вибір критеріїв ефективності. Можна отримати найкращі комбінації відповідно до SROCC, PCC або середнім значенням SROCC та PCC. Вони були різними, хоча більшість елементарних показників були однаковими. Для подальшого аналізу залишилось лише п'ять показників, наведених у Таблиці 4. Як впливає з даних Таблиці 2, усі вони є одними з найкращих, і, узагальнюючи, вони добре доповнюють одне одного.

По-перше, хороші результати були отримані для стійкої комбінованої метрики для бази даних TID2013. Як бачимо, найкращі значення SROCC та PCC для розглянутих підмножин становлять приблизно 0,95, тобто достатньо більше, ніж для будь-якої елементарної метрики, що використовується в наборі. Тим часом бажано мати метрики, які добре працюють для різних баз даних; тому рекомендується перевірка для різних баз даних. Як зазначалося на початку, лише декілька баз даних зображень можуть бути частково використані для аналізу зображень дистанційного зондування. Однією з таких баз даних, запропонованих нещодавно, є KADID10k. Вона містить меншу кількість зображень, що містять спотворення, перелічені в Таблиці 2, які використовувались для аналізу, також розділені на дві групи (див. Табл. 5).

Аналіз даних для бази даних KADID10k показує, що елементарні метрики ADM та CVSSI працюють досить добре. Тим часом результати для запропонованої комбінованої метрики є або найкращими, або близькими до найкращих для елементарних метрик (див. дані, позначені жирним шрифтом). Цікаво також, що запропонована комбінована метрика дає хороші результати для всіх типів спотворень у базі даних KADID10k.

Запропонована метрика також повинна мати можливість працювати в реальних практичних умовах, коли на зображенні одночасно можуть бути присутніми декілька типів спотворень. Для моделювання таких ситуацій використовуються бази даних зображень MDID та LIVE з множинними спотвореннями (LIVE MD) [47]. MDID включає зображення з чотирма



спотвореннями (розмиття, адитивний шум, зміна контрасту, стиснення JPEG та JPEG2000), випадково накладене на зображення відповідно до заданого алгоритму [42]. LIVE MD складається з двох пар спотворень (розмиття з подальшим JPEG та розмиття з подальшим шумом). Обидві бази даних важко розділити на групи спотворень. З цієї причини вони представлені в Таблиці 5 із загальними оцінками SROCC та PCC.

**Таблиця 5 – Результати для елементарних метрик та найкращої комбінованої**

Види спотворень	PSNRHA	ADM	DSI	PSIM	VSSI	$\alpha$ -trim
TID2013 Результати для SROCC						
Bci	0,8198	0,7861	0,8897	<b>0,8926</b>	0,8285	0,8825
Noise	0,9230	0,9113	0,9275	0,9189	0,9248	<b>0,9393</b>
Actual	0,9388	0,9201	0,9387	0,9309	0,9350	<b>0,9484</b>
Noise & Actual	0,9322	0,9189	0,9374	0,9303	0,9341	<b>0,9473</b>
TID2013 Результати для PCC						
Bci	0,8339	0,8021	0,9051	0,9039	0,7243	<b>0,9055</b>
Noise	0,9127	0,8916	0,9285	0,9202	0,9205	<b>0,9410</b>
Actual	0,9298	0,9012	0,9434	0,9336	0,9381	<b>0,9540</b>
Noise & Actual	0,9266	0,8998	0,9418	0,9326	0,9367	<b>0,9528</b>
KADID10k Результати для SROCC						
Bci	0,8252	0,8033	0,8852	0,8698	0,6939	<b>0,8922</b>
Noise	0,8565	0,9060	0,8988	0,8845	<b>0,9130</b>	0,9093
Actual	0,8673	<b>0,9380</b>	0,9190	0,9095	0,9335	0,9302
Noise & Actual	0,8681	<b>0,9356</b>	0,9202	0,9066	0,9324	0,9294
KADID10k Результати для PCC						
Bci	0,7687	0,7907	0,8436	0,8629	0,5261	<b>0,8790</b>
Noise	0,8463	0,8893	0,8962	0,8791	0,9082	<b>0,9097</b>
Actual	0,8458	0,9177	0,9098	0,9044	<b>0,9270</b>	0,9268
Noise & Actual	0,8524	0,9126	0,9132	0,9060	0,9266	<b>0,9278</b>
Результати для бази MDID						
Bci (SROCC)	0,7240	0,8283	0,8360	0,8733	<b>0,8615</b>	0,8576
Bci (PCC)	0,7164	0,8255	0,8121	0,8593	<b>0,8613</b>	0,8566
Результати для LIVE MD						
Bci (SROCC)	0,7111	<b>0,8821</b>	0,7764	0,8508	0,8760	0,8326
Bci (PCC)	0,7456	<b>0,8923</b>	0,8106	0,8836	0,8923	0,8652



Що стосується бази даних MDID, то метрика CVSSI працює найкращим чином, хоча запропонована комбінована метрика також працює добре. Що стосується бази даних LIVE MD, то найкращим чином працює метрика ADM. Результати для комбінованої метрики не такі хороші, але можуть бути задовільними для практики. Згідно із загальними результатами, стійке поєднання MDSI, PSNRHA, ADM, PSIM та CVSSI після лінеаризації можна вважати найкращим набором елементарних метрик.

Таким чином, проведено аналіз ефективності для 50 метрик візуальної якості з еталоном для типів спотворень, типових для зображень дистанційного зондування. З їх використанням було запропоновано стійкі лінеаризовані комбіновані метрики, засновані на альфа-урізаному середньому значенні п'яти оцінок. Також були розглянуті аспекти підгонки (фітінгу) та лінеаризації. Оптимізація шляхом пошуку найкращих наборів елементарних метрик серед доступних проведена для бази даних TID2013 з використанням критеріїв як SROCC, так і PCC. Було показано, що використання альфа-урізаного середнього, застосованого до оцінок MOS, дозволяє досягти значень SROCC та PCC близько 0,95 для підмножин TID2013 "Noise" та "Actual". Пропоноване рішення було перевірено для інших існуючих баз даних (KADID10k, MDID та LIVE MD), що приводить до SROCC та PCC приблизно 0,86.

## 4.2. Комбіновані метрики з використанням нейронних мереж

В практиці дистанційного зондування можливі різні ситуації:

– зображення здається ідеальним, тобто візуально не можна виявити погіршення якості (чіткість задовільна, шуму не видно, інших погіршень не спостерігається);

– зображення багатоканальне, і є компонентні зображення дуже високої якості та компонентні зображення досить низької якості [24, 25]; для даних ДЗ з великою кількістю компонентів, тобто гіперспектральних зображень, це можна виявити шляхом компонентної візуалізації та аналізу зображень;

– отримане зображення з самого початку є певним чином спотвореним, наприклад, завдяки принципу роботи системи формування зображення; хорошими прикладами є зображення радіолокаційних систем із синтетичною



апертурою (SAR), для яких завжди присутній спекл-шум [49].

Це означає, що якість оригінальних (первинних) зображень повинна характеризуватися кількісно за допомогою певних показників. Слід також охарактеризувати ефективність попередньої обробки зображення ДЗ (наприклад, усунення шуму або стиснення з втратами). У цьому сенсі існує кілька груп метрик (критеріїв), які можуть бути використані для цієї мети. По-перше, бувають практичні ситуації, коли можна застосовувати метрики з еталоном (див. підрозділ 1). Це трапляється, наприклад, при стисненні даних із втратами, коли метрику можна обчислити, використовуючи оригінальні та стиснені зображення [50–52]. По-друге, метрики без еталону можна використовувати, коли дані без спотворень відсутні. Зазвичай деякі параметри зображення розраховуються для обчислення метрики з відсутністю еталону. Зауважте, що є досить успішні спроби прогнозувати метрики з еталоном, не маючи при цьому еталононих зображень [53, 54]. Нарешті, існує багато метрик, які характеризують якість зображення (або ефективність обробки зображень) з точки зору якості вирішення кінцевих завдань [55–57]. Це може бути, наприклад, площа під кривою (AUC – area under the curve) [58] або ймовірності правильної класифікації [59].

Очевидно, що багато метрик є корельованими між собою. Наприклад, критерії класифікації зображень ДЗ залежать від якості вхідних даних, хоча ефективність класифікації також залежить від використовуваного набору ознак, застосованого класифікатора та використовуваного методу навчання. В цьому підрозділі зосередимось на метриках, що характеризують якість первинних зображень або зображень після попередньої обробки як-то фільтрація або стиснення з втратами, фокусуючись на метриках з еталоном, і, зокрема, на метриках візуальної якості.

Звичайні показники, такі як середньоквадратична похибка (MSE mean square error) або пікове відношення сигнал/шум (PSNR), все ще широко використовуються для аналізу зображень ДЗ або оцінки ефективності їх обробки [57, 60–62]. Тим часом існує очевидна тенденція застосовувати візуальні показники якості [36, 37, 61, 63–66]. Є кілька робіт, що використовують SSIM [67], який є, мабуть, найдавнішим - крім UQI [68], будучи його безпосереднім попередником - метрикою візуальної якості [61, 63, 64]; деякі інші показники візуальної якості були розроблені та протестовані нещодавно [65–68] для конкретних додатків, таких як посилення різкості зображення та виявлення об'єктів. Однак кількість робіт, де використовуються показники візуальної



якості, все ще обмежена [61–66, 69–71].

Тим не менше, існує кілька причин застосовувати показники візуальної якості. Відомо, що система людського зору приділяє першочергову увагу чіткості зображення [72–75], будучи високо корельованою як з метриками якості зору, заснованими на HVS, так і з багатьма завданнями обробки зображень, такими як фільтрація, стиснення з втратами, виявлення країв, сегментація та класифікація. Отже, цілком розумно знати, які відомі метрики можна застосувати, які найкращі серед них, і чи можна покращити показники цих метрик. Подібні проблеми вже частково вирішені в мультимедійних додатках, де: а) запропоновано багато HVS-метрик [72–75]; б) запропоновані методології їх тестування (верифікації) [72–75] в) створені бази даних для верифікації метрик [40–42, 72–76]; г) зроблені попередні висновки [72–76], які дозволяють знайти хороші показники для певного виду (типів) спотворень; д) висунуто шляхи підвищення ефективності метрик, включаючи проектування комбінованих метрик [77–80] або метрик на основі нейронної мережі (neural network - NN) (деякі приклади наведені в [77–80]).

В розділі 4.1 пояснено головні проблеми розробки метрик візуальної якості для такого застосування, як дистанційне зондування. Також наведено аргументи на користь використання бази TID2013 та наборів зображень із певними типами спотворень. Втім, метод об'єднання кількох елементарних метрик за допомогою попередньої лінеаризації та стійких оцінок може бути не найбільш ефективним із можливих. Тому нами було також досліджено інший (дуже популярний у даний час) підхід на основі нейромереж.

В останні роки нейронні мережі (NN) продемонстрували дуже високий потенціал у вирішенні багатьох завдань, пов'язаних з обробкою зображень. Їх використання часто трактується як засіб для отримання переваг у покращенні розробки та ефективності. Отже, коротко розглянемо особливості та можливості використання NN для нашого застосування, тобто при розробці нових, більш потужних метрик з еталоном для зображень із вищезгаданими типами спотворень. Вимоги до таких метрик наводяться нижче.

Хороша метрика на основі NN повинна забезпечувати певну перевагу в роботі порівняно з елементарними метриками. Оскільки ми розглядаємо SROCC як головний кількісний критерій, його слід значно покращити порівняно з уже досягнутими значеннями 0,93 ... 0,94. Оскільки максимальне значення SROCC дорівнює одиниці, його покращення на 0,02... 0,03 можна вважати достатнім.





Іншими важливими аспектами є вхідні параметри та структура NN. Оскільки типовою вимогою до метрики з еталоном є досить швидка робота, вхідні параметри слід обчислювати легко і швидко. Звичайно, їх обчислення можна зробити паралельно або якось прискорити, але в будь-якому випадку жоден з вхідних параметрів не повинен бути занадто складним. Структура використовуваної NN повинна бути, наскільки можливо, також простою. Потрібна менша кількість прихованих шарів і менша кількість нейронів у них без втрати ефективності. Менша кількість вхідних параметрів також може бути вигідною у плані спрощення та швидкодії.

Аналіз існуючих рішень показує наступне:

– нейронні мережі вже використовувались при проектуванні метрик якості з еталоном (див., наприклад, [81–84]); метрика [81] використовує ознаки з еталонних та спотворених зображень та використовує глибоке навчання у розробці метрик, забезпечуючи SROCC = 0,94 для всіх типів спотворень у TID2013; Е. Прашнані та ін. [82] трохи покращили результати роботи [81] завдяки використанню нової системи парного навчання; Сео та ін. [83] досяг SROCC = 0,961 за допомогою глибокого навчання;

– можуть існувати різні структури NN (незважаючи на популярність згорткових мереж, стандартні багат шарові мережі все ще можуть бути досить ефективними) і різні набори вхідних параметрів (можуть використовуватися як деякі функції, так і елементарні метрики).

Маючи це на увазі, наша ідея полягає у використанні набору елементарних метрик якості як вхідних даних та застосуванні NN з досить простою структурою для вирішення нашого завдання - отримати комбіновану метрику (або кілька комбінованих метрик) з ефективністю, суттєво кращою, ніж для найкращої елементарної метрики. Тоді виникає набір конкретних завдань, які потрібно вирішити, а саме:

- скільки елементарних показників слід використовувати?
- які елементарні показники слід використовувати?
- яку структуру нейромережі слід обрати та як оптимізувати її параметри?
- чи потрібні деякі операції попередньої обробки вхідних даних?
- яким повинен бути вихід NN та його властивості?

Щодо останнього питання, то оскільки ми плануємо використовувати TID2013, слід нагадати, що якість зображень у цій базі даних характеризується



MOS. Основні властивості MOS в TID2013 визначаються методологією експериментів, проведених спостерігачами. Потенційно могло бути, що MOS може бути від 0 до 9, але, як результат експериментів, MOS коливається в межах від 0,24 до 7,21 [85]. Більше того, аналіз MOS та якості зображення [85] показав, що щодо MOS можливі чотири градації якості зображення:

- а) відмінна якість ( $MOS > 6,05$ );
- б) гарна якість ( $5,25 < MOS \leq 6,05$ );
- в) середня якість ( $3,94 < MOS \leq 5,25$ );
- г) погана якість ( $MOS \leq 3,94$ ).

Ця "класифікація" трохи суб'єктивна, тому потрібні деякі пояснення. Зображення вважаються такими, що мають відмінну якість, якщо спотворення на них неможливо візуально помітити. Для зображень з гарною якістю значення MOS мають ранги від 201 до 1000, а спотворення можна помітити при ретельному візуальному огляді. Якщо значення MOS мають рейтинг від 1001 до 2000, якість зображення класифікується як середнє (спотворення видно, але вони не дратують). Якість інших зображень умовно класифікується як погана - спотворення здебільшого дратують.

Рисунок 2 ілюструє приклади спотворених зображень для одного й того ж еталонного зображення (№19 у TID2013), яке має нейтральний вміст і має середню складність. Для них також представлені значення MOS та три елементарні показники. Зображення на рисунку 2,а) відповідає першій групі (відмінна якість), і виявити спотворення справді важко. Зображення на рисунку 2,б) належить до другої групи, і спотворення видно, особливо в однорідних областях зображення. Зображення на рисунку 2,в) є хорошим представником третьої групи зображень, для яких спотворення очевидні, але вони не дратують. Нарешті, зображення на рисунку 2,г) є прикладом неякісного зображення. Представлені значення метрик показують, наскільки вони відповідають погіршенню якості та можуть характеризувати рівень спотворень.

Передостаннє запитання стосується попередньої обробки вхідних даних. Відомо, що в теорії нейромереж часто рекомендується здійснити деяку попередню нормалізацію вхідних даних (ознак), якщо вони мають різні діапазони варіювання. Це справедливо для елементарних метрик, як, наприклад, PSNR і, наприклад, PSNR-HVS-M, обидва виражені в дБ, можуть варіюватися в широких межах (від 10 дБ до 60 дБ), але для спотворених зображень у TID2013 вони варіюються у вузьких межах.



PSNR, який використовується для встановлення п'яти рівнів спотворень, має п'ять значень, що приблизно дорівнюють 21, 24, 27, 30 та 33 дБ, хоча насправді значення цієї метрики для зображень у TID2013 коливаються приблизно від 13 дБ до  $\approx 41$  дБ. PSNR-HVS-M варіюється від 14 дБ до 59 дБ, що в основному відповідає “робочим межах”, починаючи від дуже дратівливих спотворень і закінчуючи практично ідеальною якістю (невидимі спотворення). MDSI варіюється від 0,1 до 0,55 для зображень у TID2013, де більші значення відповідають нижчій якості зору.



а)



б)



в)



г)

а) відмінне (тип спотворення №2 - адитивний шум у кольорових компонентах) -

MOS = 6,0811, PSNR = 43,6260 дБ, PSNR<sub>H</sub>A = 39,8060 дБ, MDSI = 0,1556;

б) гарне (тип спотворення №5 - високочастотний шум) - MOS = 5,9737, PSNR =

37,7730 дБ, PSNR<sub>H</sub>A = 42,4980 дБ, MDSI = 0,1512;

в) середнє (тип спотворення №3 - просторово корельований шум) - MOS =

4,6667, PSNR = 37,8280 дБ, PSNR<sub>H</sub>A = 33,4380 дБ, MDSI = 0,2683;

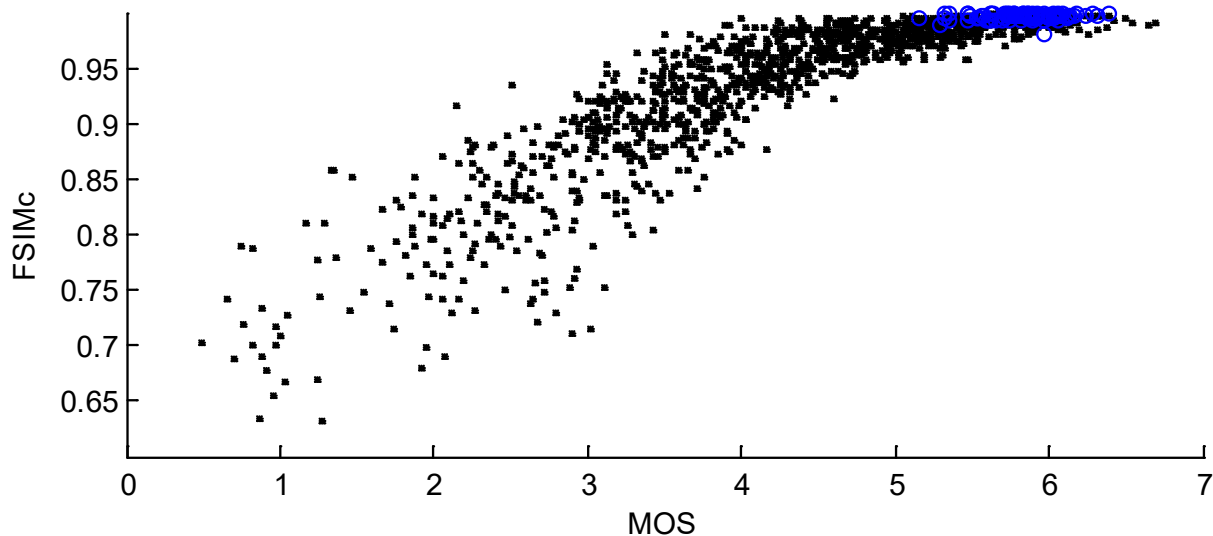
г) погана (тип спотворення №6 - імпульсний шум) - MOS = 3,8844, PSNR =

31,2660 дБ, PSNR<sub>H</sub>A = 31,2890 дБ, MDSI = 0,2920

**Рисунок 2 – Приклади тестового зображення з різною якістю**



Деякі інші показники, такі як SSIM, MS-SSIM та FSIM, варіюються в межах від 0 до 1, де одиниця відповідає ідеальній якості. Насправді більшість значень цих показників зосереджено у верхній третині цього інтервалу (див. графік розсіювання на рисунку 3 для кольорової версії метрики FSIM, що називається FSIMc). Спостерігається очевидна загальна тенденція збільшення FSIMc, коли MOS стає більшим. Тим часом можна також помітити два важливі явища. По-перше, існує певне варіювання (дисперсність) значень метрики для тієї самої MOS. По-друге, залежність FSIMc від MOS (або MOS від FSIMc) нелінійна.



**Рисунок 3 – Скатерограма FSIMc від MOS (сині позначки відповідають зображенням, спотворення яких спостерігачами не виявлено)**

У цьому сенсі лінеаризація (фітінг) часто використовується для отримання не тільки високих значень SROCC, але й звичайного коефіцієнта кореляції Пірсона. Тоді можливі дві гіпотези. Перша з них полягає в тому, що неймережі, будучи нелінійним і здатним адаптуватися до особливостей вхідних даних, буде «керувати» цією нелінійністю залежності вхід-вихід «самостійно» (позначимо цю гіпотезу як H1). Друга гіпотеза (H2) полягає в тому, що попередня обробка у вигляді фітінгу може бути корисною для подальшого поліпшення комбінованих метрик (їх оптимізації).

Щодо фітінгу, необхідного для реалізації H2, існує кілька загальноприйнятих варіантів. Одним з них є застосування функції виду  $y(x) = a \cdot xb + c$ , де  $a$ ,  $b$  та  $c$  - регульовані параметри. Результати підгонки можна охарактеризувати середньоквадратичним відхиленням (RMSE) точок діаграми розсіювання (скатерограми) після фітінгу відносно встановленої кривої (чим



менше RMSE, тим краще). Результати, отримані для елементарних метрик, розглянутих вище (у табл. 1), наведені в таблиці 6. Як бачимо, найкращий результат отриманий для метрики MDSI (він дорівнює 0,3945). Результати для метрик, які є одними з найкращих за даними SROCC (наприклад, CVSSI, MCSD, PSNRHA, GMSD, див. дані в таблиці 1), майже однаково хороші.

Іншою придатною моделлю, яку можна застосувати, є, наприклад, поліноміальна Poly2:  $y(x) = p_1 \cdot x^2 + p_2 \cdot x + p_3$ , де  $p_1$ ,  $p_2$  та  $p_3$  - це параметри, які слід коригувати (варіювати) для отримання найкращої підгонки. Отримані результати дуже схожі. Наприклад, для MDSI спостерігається мінімальний RMSE, що дорівнює 0,3951. Для PSNR-HA досягається  $RMSE = 0,4013$ , тобто дещо кращий, ніж у таблиці 6. Тому докладні дані, отримані для підгонки Poly2, не представлені для економії місця, хоча результати, кращі для першої моделі, позначені жирним шрифтом у таблиці 6. Тоді при розробці комбінованих метрик на основі нейромереж можливі три варіанти: 1) використовувати першу модель (позначимо її PFF) для всіх елементарних метрик, 2) застосувати Poly2 для всіх елементарних метрик; 3) вибрати найкращий варіант для кожного елементарного показника та застосувати його.

Наступним розглянутим питанням є те, які структури та параметри нейромереж можуть бути обрані та оптимізовані. Як було зазначено вище, ми розглядаємо звичайні структури нейромереж. Причина, по якій не розглядається підхід на основі глибокого навчання, полягає в тому, що потрібен великий навчальний набір даних, і ми не маємо їх у своєму розпорядженні.

Тоді простим вибором є застосування багатошарових нейронних мереж, для яких зрозумілі багато аспекти проектування та навчання. Це зменшує складність нейронної моделі, дозволяючи добре узагальнити дані та запобігаючи проблемам перетренованості. Зокрема, кількість нейронів у вхідному шарі має бути рівною кількості використовуваних елементарних метрик. Щодо цього числа, можливі різні варіанти, включаючи використання всіх 50 розглянутих елементарних метрик. Також можна обмежитися найкращими метриками з таблиці 1 та застосувати елементарні метрики, які мають певні властивості, наприклад, метрики, які мають  $SROCC > 0,9$  для розглянутих типів спотворень (таких 22 метрики),  $SROCC > 0,92$  (14 метрики), або  $SROCC > 0,93$  (7 метрик).



Таблиця 6 – Характеристики фітінгу для елементарних метрик

Метрика	RMSE	a	b	c
PSNR	0,6830	-126,636	-0,0554	109,1416
<b>MSE</b>	<b>0,6811</b>	<b>-2,4997</b>	<b>0,1690</b>	<b>9,3328</b>
WSNR [89]	0,5216	103,726	0,0372	-113,987
NQM [96]	0,6037	0,8643	0,5904	-1,3442
<b>UQI [40]</b>	<b>0,8393</b>	<b>6,5341</b>	<b>0,4358</b>	<b>-0,8949</b>
SSIM [39]	0,7645	3,2558	4,5929	2,8783
<b>MS-SSIM [64]</b>	<b>0,5536</b>	<b>3,5705</b>	<b>12,6686</b>	<b>2,2151</b>
<b>IFC [77]</b>	<b>0,7404</b>	<b>30,6531</b>	<b>0,0374</b>	<b>-27,7813</b>
<b>VIF [71]</b>	<b>0,5600</b>	<b>4,9142</b>	<b>0,6852</b>	<b>1,1082</b>
VIFP [71]	0,6514	11,1695	0,2507	-4,6922
<b>MSVD [74]</b>	<b>1,2091</b>	<b>463,362</b>	<b>-1,7245</b>	<b>4,0580</b>
<b>QILV [101]</b>	<b>0,8924</b>	<b>2,1258</b>	<b>12,8721</b>	<b>3,2304</b>
VSNR [93]	0,5505	-40,9961	-0,1030	33,6452
PSNRHVS [82]	0,4176	-54,0574	-0,4961	14,4993
PSNRHVSM [63]	0,4472	-78,8179	-0,7655	9,8824
<b>CWSSIM [99]</b>	<b>0,7294</b>	<b>2,4648</b>	<b>1858,3573</b>	<b>3,0546</b>
RFSIM [76]	0,5086	4,9998	0,8521	1,0324
PSNRHAy [69]	0,4114	-79,7808	-0,7054	11,7318
PSNRHMAy [69]	0,4332	-120,791	-0,9339	9,0571
PSNRHA [69]	0,4057	-271,292	-1,1468	9,8101
PSNRHMA [69]	0,4296	-437,562	-1,3364	8,5281
<b>FSIM [75]</b>	<b>0,4813</b>	<b>3,6227</b>	<b>13,3796</b>	<b>2,1510</b>
<b>FSIMc [75]</b>	<b>0,4699</b>	<b>3,7012</b>	<b>11,1900</b>	<b>2,1265</b>
<b>IWSSIM [91]</b>	<b>0,5336</b>	<b>3,4602</b>	<b>11,7455</b>	<b>2,2851</b>
ADM [86]	0,5133	5,2079	1,0584	1,1913
<b>GSM [95]</b>	<b>0,6085</b>	<b>3,8941</b>	<b>62,8385</b>	<b>2,0375</b>
<b>IGM [83]</b>	<b>0,4154</b>	<b>4,8469</b>	<b>4,6108</b>	<b>0,3577</b>
<b>SR-SIM [84]</b>	<b>0,4485</b>	<b>3,6224</b>	<b>26,1269</b>	<b>2,1720</b>
<b>SFF [78]</b>	<b>0,5037</b>	<b>4,3154</b>	<b>17,0200</b>	<b>1,4609</b>
<b>GMSD [81]</b>	<b>0,4201</b>	<b>-10,7552</b>	<b>0,5948</b>	<b>6,2403</b>
<b>ESSIM [94]</b>	<b>0,5862</b>	<b>4,1907</b>	<b>794,155</b>	<b>1,7213</b>
WASH [102]	1,1077	-3,9034	-0,8187	9,1013
<b>VSI [68]</b>	<b>0,4577</b>	<b>3,5501</b>	<b>37,8918</b>	<b>2,3297</b>
IQM2 [88]	0,4716	3,6621	1,2572	2,1072
<b>DSS [90]</b>	<b>0,5109</b>	<b>3,3472</b>	<b>3,2571</b>	<b>2,2928</b>
<b>ADD_GSIM [87]</b>	<b>0,4544</b>	<b>3,7670</b>	<b>155,109</b>	<b>1,9912</b>
<b>ADD_SSIM [87]</b>	<b>0,4705</b>	<b>3,9271</b>	<b>142,763</b>	<b>1,8464</b>
<b>MCSD [80]</b>	<b>0,4123</b>	<b>-10,6097</b>	<b>0,5096</b>	<b>6,4149</b>
<b>MDSI [66]</b>	<b>0,3945</b>	<b>-17,9432</b>	<b>1,6790</b>	<b>6,7208</b>
<b>UNIQUE [98]</b>	<b>0,6600</b>	<b>4,3775</b>	<b>0,4905</b>	<b>1,3221</b>
<b>MSUNIQUE [98]</b>	<b>0,6478</b>	<b>4,3483</b>	<b>0,6346</b>	<b>1,4211</b>
<b>PSIM [67]</b>	<b>0,4144</b>	<b>3,9219</b>	<b>583,789</b>	<b>1,9809</b>
<b>CVSSI [79]</b>	<b>0,4137</b>	<b>-10,9029</b>	<b>0,5469</b>	<b>6,3852</b>
PSNRHMAm [70]	0,4086	-258,9957	-1,1396	9,7006
<b>DSI [100]</b>	<b>0,7916</b>	<b>-2,4387</b>	<b>0,1568</b>	<b>7,6098</b>
<b>CSSIM [92]</b>	<b>0,5238</b>	<b>3,3778</b>	<b>19,8756</b>	<b>2,2658</b>
<b>CSSIM4 [92]</b>	<b>0,6121</b>	<b>2,8845</b>	<b>72,9625</b>	<b>2,6397</b>
<b>SSIM4 [92]</b>	<b>0,5315</b>	<b>3,2548</b>	<b>9,2413</b>	<b>2,3957</b>
<b>HaarPSI [85]</b>	<b>0,4347</b>	<b>4,5283</b>	<b>2,1854</b>	<b>1,4475</b>
<b>RVSIM [97]</b>	<b>0,5678</b>	<b>5,5441</b>	<b>1,1196</b>	<b>-0,0078</b>



В той же час теорія нейромереж стверджує, що доцільно застосовувати такі вхідні параметри, які можуть "додавати" або "доповнювати" інформацію один до одного, тобто не мають високої кореляції. Одним із можливих підходів є оцінка функції перехресної кореляції, щоб визначити подібність між усіма парами метрик, а після цього виключити найгірші у висококорельованих парах.

Замість цього в цій роботі застосовується дуже корисний підхід, який називається регуляризацією Лассо [86], для вибору найбільш унікальних метрик для NN. При машинному навчанні регуляризації Лассо використовуються для введення додаткових обмежень у модель та зменшення проблеми перенавчання. Ключовою особливістю Лассо є те, що цей метод може вводити нульові ваги для "шумних" та найменш важливих даних. Для завдання комбінації метрик це означає, що регуляризація Лассо може визначити елементарні метрики, які є найменш корисними для комбінування, і залишити інші («найбільш інформативні»). Ще однією перевагою є наявність функції Лассо в MATLAB, яка з різними пороговими значеннями оцінює нульові коефіцієнти для метрик, які можна виключити. Відповідно до цього було визначено кількість таких ненульових значень (NNZ) для кожної метрики, приймаючи такі умови: а)  $NNZ > 20$  (із 100 значень); б)  $NNZ > 30$ ; в)  $NNZ > 40$ ; г)  $NNZ > 50$ ; д)  $NNZ > 60$ .

Оскільки вихідом нейромережі є комбінована метрика, то у вихідному шарі присутній один нейрон, але при цьому кількість прихованих шарів і кількість нейронів у кожному прихованому шарі можуть бути різними. Проаналізовано два варіанти - два та чотири прихованих шари. Крім того, були розглянуті два варіанти кількості нейронів у прихованих шарах, а саме однакова кількість нейронів у кожному прихованому шарі та вдвічі менша кількість нейронів у кожному наступному прихованому шарі. Як функція активації в прихованих шарах використовується гіперболічна дотична сигмоїдна функція, яка також забезпечує нормалізацію в діапазоні  $(-1,1)$ , тоді як лінійна функція використовується для вихідного шару.

Приклад структури NN представлений на рисунку 4. У цьому випадку є 9 входів, і кількість нейронів у прихованих шарах поступово зменшується.

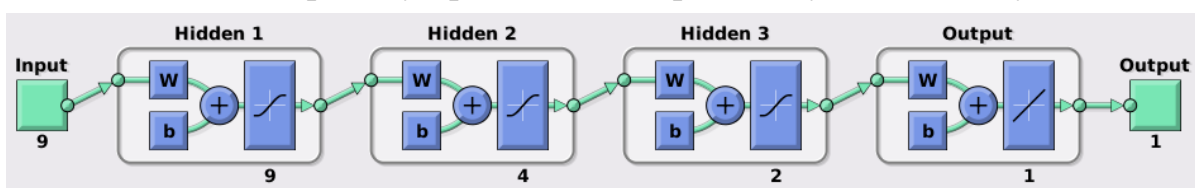


Рисунок 4 – Приклад структури нейронної мережі



На додаток до структури NN, на її ефективність можуть впливати деякі інші фактори, такі як методологія навчання та тестування NN, стабільність результатів навчання, кількість епох тощо. Через обмежений набір 1625 зображень для навчання NN та тестування, наше завдання ускладнюється. Відповідно до традиційної методології, наявний набір має бути розділений на навчальний та перевірочний у певній пропорції. У проведених експериментах 70% зображень було використано для тренувань, а решта зображень для перевірки. Оскільки розподіл зображень на набори є випадковим, результати навчання та перевірки також можуть бути випадковими. Щоб частково обійти цю невизначеність, найкращі дані представлені нижче для кожної версії навченої NN (що забезпечує найбільший SROCC на етапі навчання) даної структури.

Щодо етапу навчання - кожна нейромережа пройшла навчання, щоб забезпечити якомога більше значення SROCC. Очевидно, можливі також деякі інші навчальні стратегії. Зокрема, можна використовувати коефіцієнт кореляції Пірсона (Pearson correlation coefficient - PCC) для максимізації. Тим не менше, результати додаткового використання PCC для навчання в даному розділі не розглядаються.

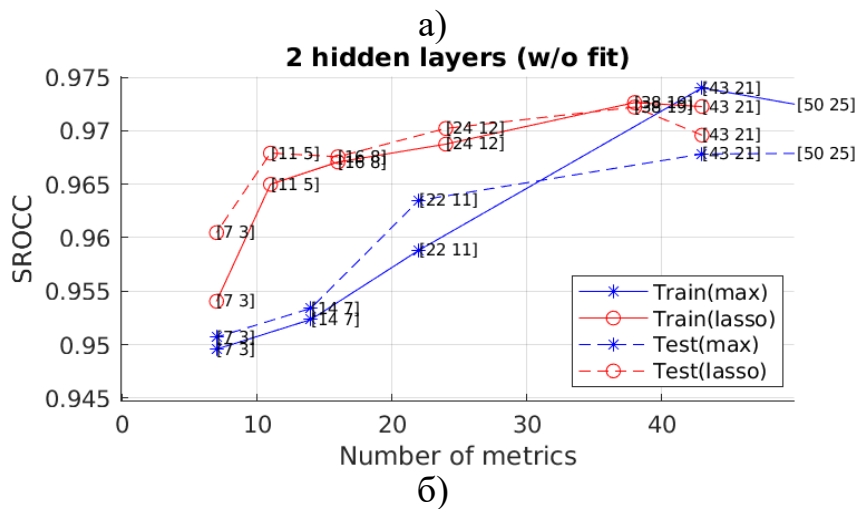
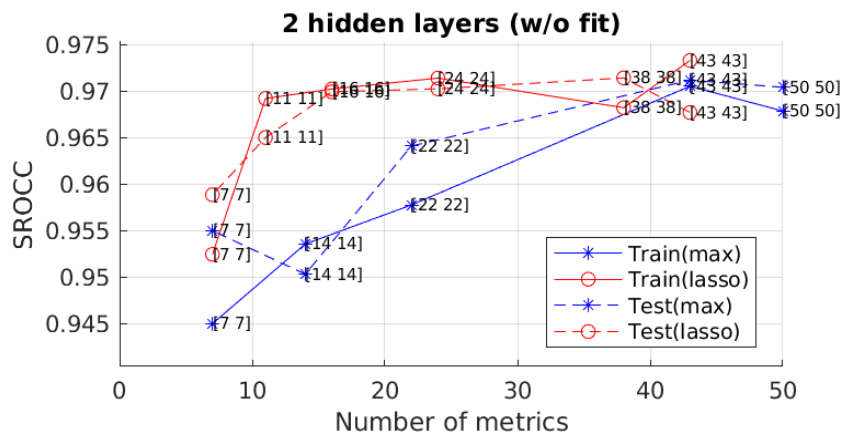
Для глибшого розуміння того, які структури нейромереж були проаналізовані та які параметри використані, основні характеристики представлені в таблиці 7, де в кожному випадку надається кількість вхідних елементарних метрик. Як бачимо, існує досить багато можливих структур нейронних мереж.

Основними критеріями навчання та верифікації нейромереж є значення SROCC. Були проаналізовані чотири SROCC: SROCCtrain (Max), SROCCtrain (Lasso), SROCCtest (Max) і SROCCtest (Lasso), які відповідають навчальним і тестовим (верифікаційним) даним із використанням максимальної та визначеної Лассо кількості входів. Починаючи від нейромережі з двома прихованими шарами, де попередній фітінг не використовується, отримані результати представлені на рисунку 5. Числа біля кожної точки на представлених графіках показують кількість нейронів у прихованих шарах. Дані на рисунку 5,а) стосуються випадку, коли кількість нейронів у прихованих шарах однакова, тоді як графіки на рисунку 5,б) відповідають ситуації, коли кількість нейронів поступово зменшується.



**Таблиця 7 – Кількість елементарних метрик, що використовуються як вхідні дані для різних конфігурацій нейромереж в умовах різного застосування та обмежень, накладених на набір елементарних метрик**

Правило попередньої обробки	Без фітінгу	PFF	Краща з PPF та Poly2
Всі 50	50	50	50
3 SROCC > 0.8	43	43	43
3 SROCC > 0.9	22	22	22
3 SROCC > 0.92	14	14	14
3 SROCC > 0.93	7	7	7
3 Lasso (NNZ > 20)	43	39	41
3 Lasso (NNZ > 30)	38	25	29
3 Lasso (NNZ > 40)	24	20	21
3 Lasso (NNZ > 50)	16	12	11
3 Lasso (NNZ > 60)	11	9	6
3 Lasso (NNZ > 70)	7	6	6



а) однакова кількість нейронів у прихованих шарах;  
 б) вдвічі менша кількість нейронів у прихованих шарах.

**Рисунок 5 – SROCC для NN з двома прихованими шарами з різною кількістю входів**



Аналіз даних показує наступне:

– якщо Лассо не використовується, збільшення кількості входів призводить до загальної тенденції збільшення як  $SROCC_{train}$ , так і  $SROCC_{test}$ ; тим часом, для більш, ніж 40 входів, покращення не спостерігається;

– якщо застосовується регуляризація Лассо, результати не настільки хороші, якщо кількість входів ( $N_{inp}$ ) менше 20; але якщо  $N_{inp}$  перевищує 20, то ефективність практично не залежить від  $N_{inp}$ ; це означає, що метод Лассо дозволяє спростити структуру нейромережі, мінімізуючи кількість входів та нейронів в інших шарах;

– найкращі (найбільші) значення  $SROCC$  перевищують 0,97, що демонструє, що завдяки використанню нейромереж та оптимізації їх параметрів досягається достатнє покращення (порівняно з найкращою елементарною метрикою);

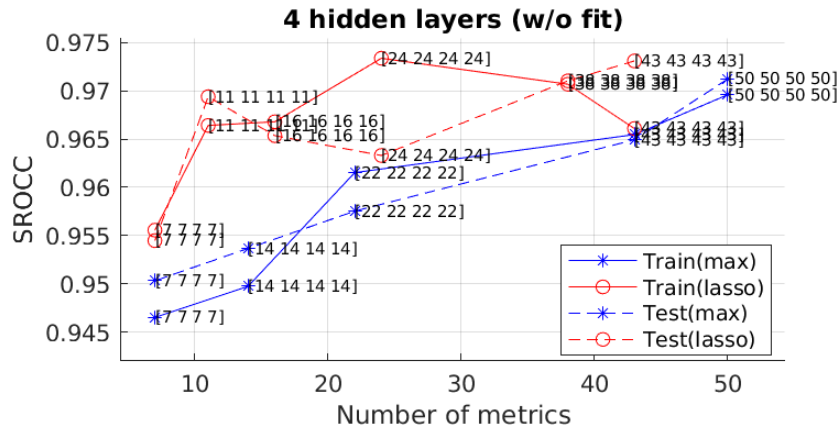
– значення  $SROCC_{train}$  та  $SROCC_{test}$  практично однакові для кожної конфігурації аналізованої нейромережі, тому результати навчання можна вважати стабільними;

– не виявлено суттєвої різниці в результатах для однакової або нерівної кількості нейронів у прихованих шарах.

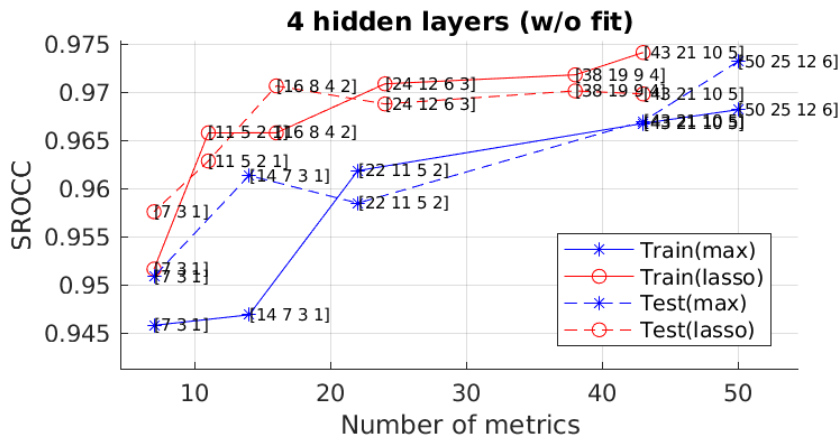
Щодо іншої кількості прихованих шарів, а саме чотирьох, то отримані графіки наведені на рисунку 6. Аналіз отриманих даних дозволяє зробити два основних висновки. По-перше, немає очевидних переваг у порівнянні з випадком використання нейромереж з двома прихованими шарами. По-друге, всі інші висновки, наведені вище, щодо  $SROCC_{train}$  та  $SROCC_{test}$ , впливу регуляризації  $N_{inp}$  та Лассо однакові, тобто в якості вхідних даних доцільно застосовувати обмежену кількість, наприклад, 24 елементарні метрики, визначені Лассо.

Наступне запитання, на яке потрібно відповісти: "чи допомагає попередній фітинг?" Відповідь на нього представлена на рисунку 7 із двома наборами графіків. Обидва набори отримані для нейромережі з двома прихованими шарами та однаковою кількістю нейронів у них. Графіки на рисунку 7,а) отримані для PPF, а на малюнку 7,б) - для найкращого фітингу. Порівняння графіків, що відповідають один одному на рисунках 7, показує, що немає достатньої різниці у виборі фітингу. Більше того, порівняння з відповідними графіками на рисунку 5,а) вказує на те, що попередній фітинг не забезпечує достатнього поліпшення характеристик у порівнянні з випадками, коли він не використовується. Це означає, що навчені нейромережі забезпечують цю

попередню обробку самостійно.



a)

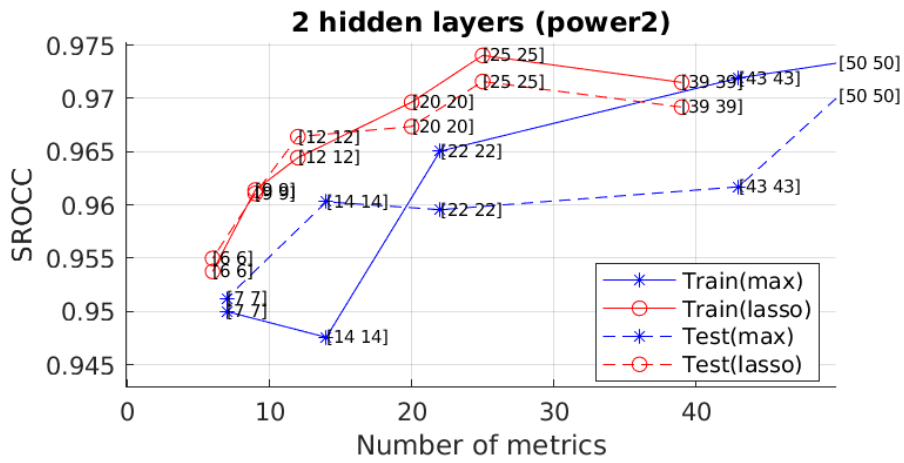


б)

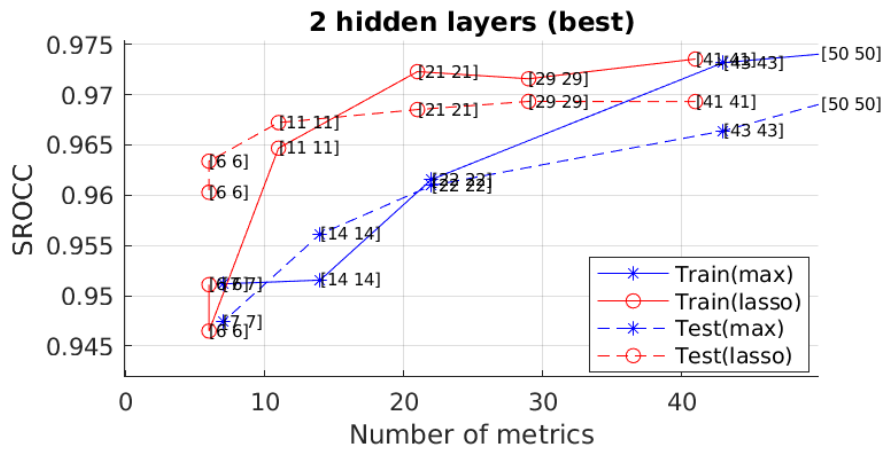
- а) однакова кількість нейронів у прихованих шарах;
- б) вдвічі менша кількість нейронів у прихованих шарах.

**Рисунок 6 – Значення SROCC для нейромереж з чотирма прихованими шарами з різною кількістю входів**

Інші висновки, які випливають з аналізу залежностей на рисунку 7, практично такі ж, як і раніше. Метод Лассо гарантує, що ефективність, близьку до оптимальної, може бути забезпечена для  $N_{inp}$  трохи більше 20. Максимально досягнуті значення  $SROCC$  перевищують 0,97 і менше 0,975. Також були проаналізовані результати для чотирьох прихованих шарів та нерівного числа нейронів у прихованих шарах, і найкращі результати виявились практично на одному рівні.



а)



б)

а) апроксимація PPF;

б) найкращий фітінг (з двох варіантів).

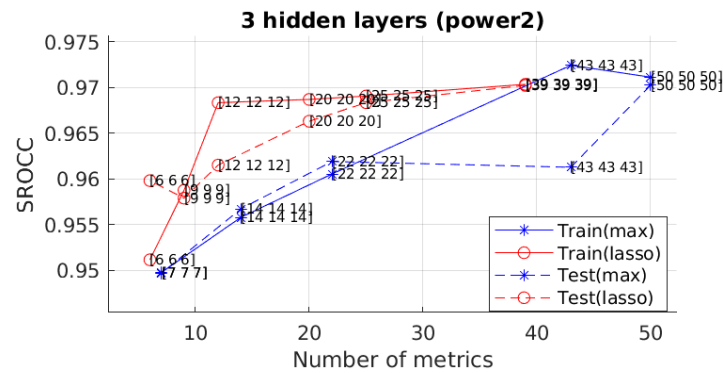
**Рисунок 7 – Значення SROCC для неймереж з двома прихованими шарами з різною кількістю входів та рівною кількістю нейронів у прихованих шарах з двома способами фітінгу**

Щоб довести, що кількість прихованих шарів не має суттєвого впливу на характеристики комбінованих метрик на основі неймереж, на рисунку 8 наведені характеристики для неймереж з трьома та п'ятьма прихованими шарами, які використовували PPF для фітінгу та рівну кількість нейронів у всіх шарах. Аналіз показує, що максимально досягнуті значення SROCC навіть менші, ніж для неймереж з двома прихованими шарами.

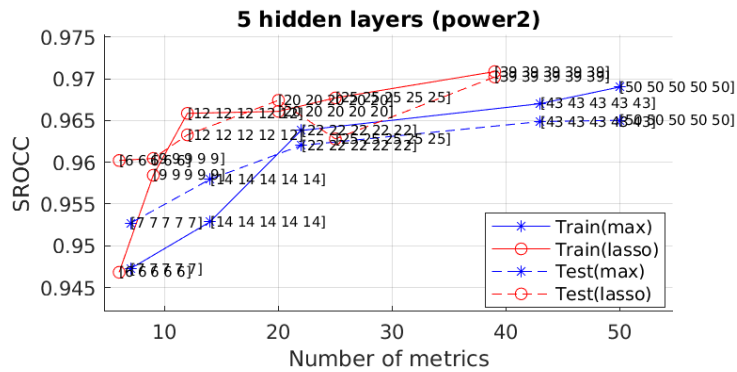
Завершуючи аналіз на основі SROCC, можна констатувати наступне:

- існує багато конфігурацій неймереж, які забезпечують приблизно однаковий SROCC;
- приймаючи до уваги бажану простоту, рекомендується використовувати

нейромережу з двома прихованими шарами без фітінгу та числом входів близько 20.



а)



б)

- а) з трьома прихованими шарами;
- б) з п'ятьма прихованими шарами.

**Рисунок 8 – Значення SROCC для нейромереж з однаковою кількістю нейронів у прихованих шарах для фітінгу PPF**

Є кілька позитивних результатів розробки з використанням Лассо. Вони стають очевидними з аналізу даних, представлених у таблиці 8. Можна помітити, що існує кілька хороших конфігурацій нейромереж, які забезпечують SROCC приблизно 0,97 для кількості входів близько 20. Разом з цим, ці нейромережі забезпечують значення RMSE, які значно менші, ніж для найкращої елементарної метрики після лінеаризації (див. дані в таблиці 2, де найкращі значення більші за 0,39). Крім того, значення коефіцієнта кореляції Пірсона (PCC) також великі і перевищують 0,97, що свідчить про дуже хорошу лінійність розроблених комбінованих метрик.



**Таблиця 8 – Характеристики розроблених комбінованих метрик якості на основі нейромереж**

#	NN configuration	Найкраща мережа			5 найкращих результатів		
		SROCC	RMSE	PCC	SROCC	RMSE	PCC
1	(NNZ60) 11 входів, 2 прихованих шари, однакова кількість нейронів, без фітингу (11/11)	0,9683	0,2873	0,9706	0,9658	0,2919	0,9697
2	(NNZ50) 16 входів, 2 прихованих шари, однакова кількість нейронів, без фітингу (16/16)	0,9702	0,2699	0,9741	0,9666	0,2863	0,9708
3	(NNZ40) 24 входа, 2 прихованих шари, однакова кількість нейронів, без фітингу (24/24)	0,9712	0,2603	0,9760	0,9694	0,2698	0,9742
4	12 входів, 2 прихованих шари, однакова кількість нейронів, фітинг power2 (12/12)	0,9652	0,2933	0,9694	0,9649	0,2961	0,9688
5	11 входів, 2 прихованих шари, однакова кількість нейронів, 'найкращий' фітинг (11/11)	0,9656	0,2927	0,9695	0,9648	0,2992	0,9681
6	16 входів, 2 прихованих шари, однакова кількість нейронів, без фітингу (16/8)	0,9672	0,2840	0,9713	0,9664	0,2890	0,9703
7	16 входів, 3 прихованих шари, однакова кількість нейронів, без фітингу (16/16/16)	0,9665	0,2912	0,9698	0,9654	0,2950	0,9691
8	16 входів, 4 прихованих шари, однакова кількість нейронів, без фітингу (16/16/16/16)	0,9664	0,2923	0,9696	0,9654	0,2982	0,9684
9	16 входів, 4 прихованих шари, однакова кількість нейронів, без фітингу (16/8/4/2)	0,9674	0,2775	0,9726	0,9673	0,2841	0,9713

Розрахувавши значення SROCC, RMSE та PCC, можна провести більш ретельний аналіз. Перше спостереження полягає в тому, що SROCC, RMSE та PCC у нашому випадку мають високу кореляцію. Більші SROCC та PCC відповідають меншим RMSE. Найкращі результати, згідно з усіма трьома критеріями, дає нейромережа з конфігурацією 3, хоча, враховуючи складність цієї нейромережі, конфігурація №2 також є хорошою. Кількість входів менше 16 (наприклад, 11 або 12 у конфігураціях №1, №4, №5) призводить до гірших значень розглянутих критеріїв. Використання конфігурацій нейромереж із попереднім фітингом, зменшенням кількості нейронів у прихованих шарах, більшою кількістю прихованих шарів (конфігурації ## 4–9) не сприяє покращенню порівняно з відповідними конфігураціями №1 та №2. У зв'язку з



цим проаналізуємо додатково застосування конфігурації нейромережі №2.

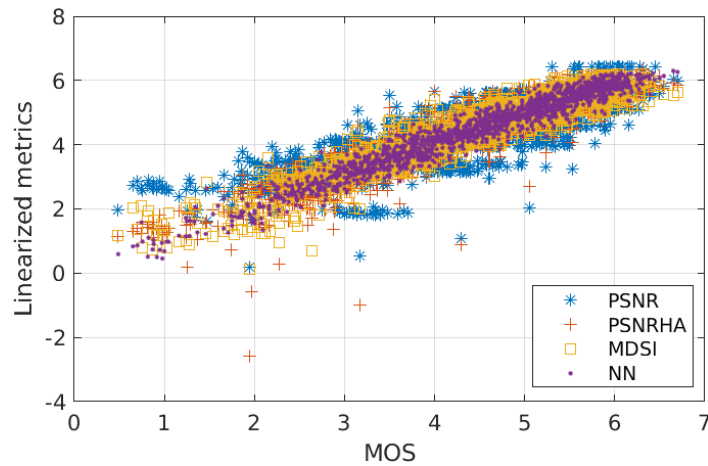
Таблиця 8 містить три стовпці, позначені заголовком «Найкраща мережа», і три стовпці, позначені «5 найкращих результатів». Раніше вже згадувалося, що результати навчання нейромережі залежать від випадкового розподілу спотворених зображень на навчальні та тестові набори. Через це для аналізу стабільності навчання ми розрахували середні значення SROCC, RMSE та PCC для п'яти найкращих результатів навчання нейромережі для кожної конфігурації. Порівняння SROCC, RMSE та PCC для перших п'яти результатів із відповідними значеннями для найкращої мережі показує, що різниця невелика. Більше того, висновки, які можна зробити в результаті аналізу цих «середніх» результатів щодо результатів роботи НС, повністю збігаються з висновками, зробленими в результаті аналізу найкращої мережі.

Аналіз ефективності метрик та їх перевірка для багатоканальних зображень ДЗ є складною проблемою. Очевидно, що найкращим рішенням може бути наявність бази даних еталонних та спотворених зображень ДЗ та MOS для кожного спотвореного зображення. Цей шлях майбутнього дослідження загалом можливий і доцільний, але вимагає значного часу та зусиль. По-перше, до експериментів повинна бути залучена велика кількість спостерігачів. По-друге, ці спостерігачі повинні володіти певними навичками в аналізі даних ДЗ, і це є основною проблемою. По-третє, набір зображень, що підлягають перегляду та оцінці, повинен бути якимось чином узгоджений у спільноті, що займається ДЗ.

Через це ми зараз можемо проводити лише деякі попередні випробування. Метою першого тесту є показати, що розроблена метрика (насправді, MOS, передбачене навченою нейромережею) дає розумні результати для конкретних зображень. Що стосується зображень на рисунку 2, то всі вони мають значення MOS, визначені в експериментах, але тепер можуть бути надані значення MOS, розраховані лінеаризацією для елементарних метрик та новою метрикою на основі NN. Отримані дані представлені в таблиці 9.

**Таблиця 9 – Істинні та спрогнозовані значення MOS**

Зображення	Істинне MOS	Оцінене MOS з використанням елементарних метрик			MOS, що оцінено за допомогою розробленої нейромережі
		PSNR	PSNRHA	MDSI	
Рисунок 2.1а)	6,08	6,4176	5,8410	5,9314	5,8329
Рисунок 2.1б)	5,97	5,5941	6,1279	5,9685	5,8697
Рисунок 2.1в)	4,67	5,6024	4,9626	4,7508	4,7202
Рисунок 2.1г)	3,87	4,5033	4,5790	4,4500	4,2479



**Рисунок 9 – Скатерограми, що ілюструють кореляцію чотирьох методів прогнозування MOS, представлених у таблиці 8, із істинними значеннями MOS (діапазон 0-7) для набору Noise & Actual бази TID2013**

Цікаву скатерограму, яку отримано для всіх розглянутих типів спотворень, показано на рисунку 9. Трохи дивним аспектом є те, що значення MOS, спрогнозовані за деякими елементарними показниками, можуть бути навіть негативними. Також видно, що метрика на основі нейромережі забезпечує високу лінійність співвідношення між істинними та спрогнозованими MOS. Також спостерігаються проблеми MDSI для малих значень MOS.

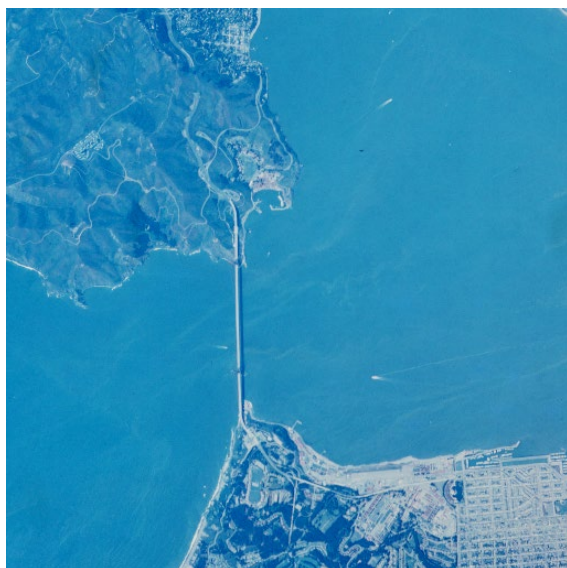
Подальші дослідження стосуються чотирьох тестових зображень, представлених на рисунку 10. Це триканальні псевдокольорові зображення, що називаються відповідно Frisco, Diego2, Diego3 та Diego4, усі розміром 512×512 пікселів, 24 біти на піксель. Причини їх вибору наступні - ці зображення мають різну складність і вони вже використовувались в деяких експериментах [87]. Зображення Frisco та Diego4 досить прості, оскільки містять великі однорідні області, тоді як зображення Diego2 має дуже складну структуру (багато дрібних деталей та текстур), а зображення Diego3 має середню складність.

Стандартною вимогою до метрик візуальній якості є монотонність, тобто монотонне збільшення або зменшення, якщо “інтенсивність” спотворення даного типу зростає. Цю властивість можна легко перевірити на її наявність для багатьох різних типів спотворень. Зображення були стиснені з використанням кодеру AGU з втратами [88], що забезпечує різну якість та ступінь стиснення (compression ratio - CR). Це було здійснено шляхом зміни кроку квантування (quantization step - QS), де більший QS відповідно пов'язаний із більшими введеними спотвореннями та гіршою візуальною якістю.





Якість (відмінна, хороша тощо) визначається відповідно до результатів у [85]. Зібрані дані представлені в таблиці 10. Очевидною тенденцією є те, що всі показники, включаючи розроблену метрику, погіршуються, якщо QS (і CR відповідно) збільшується. Значення метрики на основі нейромережі перевищують значення MOS, передбачені з елементарних метрик для тестового зображення Frisco, але менші для тестового зображення Diego2.



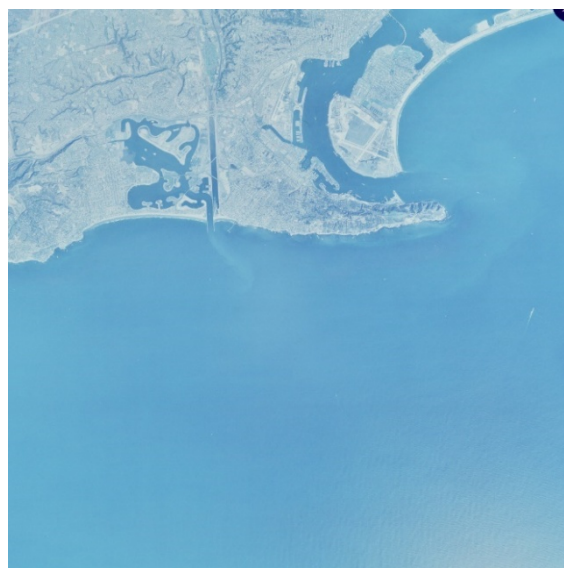
а)



б)



в)



г)

а) Frisco; б) Diego2; в) Diego3; г) Diego4.

### Рисунок 10 – Триканальні тестові зображення ДЗ

Можливо, ця властивість частково пов'язана зі складністю зображення. Однак є докази того, що розроблена метрика "поводиться" правильно. На рисунку 11 представлені стиснуті зображення із використанням  $QS = 40$ , для



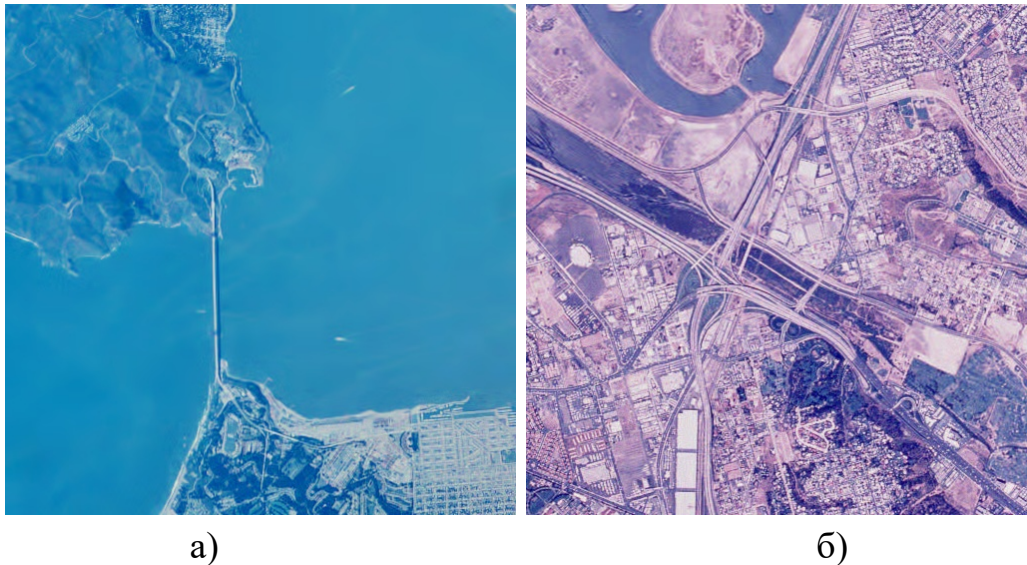
яких спотворення завжди видно. Візуальний огляд (порівняння) цих зображень із еталонними зображеннями на рисунках 10,а) та 10,б) показує, що спотворення є більш помітними для тестового зображення Diego2. Це чітко підтверджують значення PSNR-НА (36,32 дБ та 32,63 дБ відповідно - див. табл. 10). PSNR демонструє ту саму тенденцію, хоча MDSI не вказує на це. Отже, можуть зустрічатися деякі випадки, коли висновки, зроблені в результаті аналізу різних показників якості, можуть бути різними.

**Таблиця 10 – Прогнозовані значення MOS для різних метрик після фітінгу для двох тестових зображень ДЗ**

Зобр. ДЗ	QS	Якість	Значення PSNR	Прогнозоване MOS	Значення PSNRНА	Прогнозоване MOS	Значення MDSI	Прогнозоване MOS	Значення розробл. метрики
Frisco	10	excellent	43,0815	6,3461	44,8013	6,3442	0,1442	6,0263	6,5005
	20	excellent	38,7741	5,7441	40,2754	5,8940	0,1920	5,5972	6,3730
	30	good	36,5231	5,4008	37,9569	5,6184	0,2183	5,3271	6,2117
	40	good	34,9522	5,1477	36,3188	5,4009	0,2371	5,1194	6,0510
Diego2	10	excellent	42,3026	6,2421	43,8623	6,2590	0,1250	6,1743	5,6921
	20	good	36,4450	5,3885	38,0930	5,6356	0,1734	5,7735	5,4895
	30	good	32,8438	4,7884	34,8640	5,1893	0,2082	5,4338	5,2050
	40	middle	30,3845	4,3373	32,6321	4,8251	0,2347	5,1473	4,8780

Також був проведений аналіз випадку адитивного білого гаусового шуму, який був доданий до розглянутих чотирьох триканальних зображень. Розглянуто чотири значення дисперсії шуму, які відповідають чотирьом верхнім рівням спотворень, що використовуються в TID2013. Також були отримані середні значення MOS (усереднені для 25 тестових зображень у TID2013), які відповідають цим значенням дисперсії шуму. Для кожного зображення були розраховані PSNR, PSNR-НА та MDSI та визначені відповідні прогнозовані значення MOS. Також розраховано метрику на основі нейромережі. Отримані результати представлені в таблиці 11.

Аналіз показує, що значення всіх метрик погіршуються (PSNR, PSNR-НА та розрахована метрика зменшуються, а MDSI збільшується), якщо дисперсія шуму збільшується, тобто зберігається властивість монотонності. Прогнозовані значення MOS досить близькі до середніх MOS, тоді як для хорошої та середньої якості PSNR-НА, MDSI та метрика на основі нейромережі забезпечують прогнозування MOS краще, ніж PSNR. Однак для зображень низької якості ситуація протилежна.



а) Frisco; б) Diego2.

**Рисунок 11 – Тестові зображення, що стиснені з QS=40**

**Таблиця 11 – Прогнозовані значення MOS для різних метрик після фітінгу для тестових зображень ДЗ, що спотворені адитивним білим гауссовим шумом**

Зобр. ДЗ	Дисп. шуму	Якість	Значення PSNR	Прогноз. MOS	Значення PSNRHA	Прогноз. MOS	Значення MDSI	Прогноз. MOS	Розроблена метрика	Середнє MOS
Frisco	32,5	Гарна	32,99	5,40	34,33	5,62	0,21	5,38	5,86	5,67
	65	Середня	30,01	4,90	31,45	5,20	0,25	4,96	5,28	5,23
	130	Середня	27,03	4,36	28,46	4,40	0,29	4,48	4,86	4,85
	260	Погана	24,06	3,77	25,56	4,11	0,33	3,96	4,24	3,77
Diego2	32,5	Гарна	32,95	5,39	34,42	5,61	0,18	5,75	5,45	5,67
	65	Середня	29,95	4,89	31,40	5,20	0,21	5,44	5,18	5,23
	130	Середня	26,96	4,35	28,41	4,70	0,24	5,04	4,81	4,85
	260	Погана	23,98	3,75	25,63	4,13	0,29	4,54	4,32	3,77
Diego3	32,5	Гарна	32,94	5,39	34,49	5,61	0,19	5,67	5,71	5,67
	65	Середня	29,96	4,90	31,41	5,20	0,22	5,33	5,36	5,23
	130	Середня	26,97	4,35	28,42	4,70	0,25	4,92	4,93	4,85
	260	Погана	23,99	3,75	25,50	4,10	0,30	4,41	4,26	3,77
Diego4	32,5	Гарна	32,94	5,39	34,40	5,62	0,16	5,91	6,18	5,67
	65	Середня	29,97	4,90	31,40	5,25	0,19	5,62	5,84	5,23
	130	Середня	26,98	4,35	28,42	4,70	0,23	5,25	5,52	4,85
	260	Погана	23,98	3,75	25,77	4,16	0,26	4,81	5,01	3,77

Крім того, слід коротко обговорити обчислювальну ефективність. Для метрики на основі нейромережі з 16 входами розрахунок елементарних метрик PSNR, MDSI, PSNRHVS, ADM, GMSD, WASH, IQM2 та HaarPSI є дуже швидким, для розрахунку IFC та RFSIM потрібно в кілька разів більше часу, тоді як розрахунок метрик MSVD, CWSSIM, PSNRHA, IGM, DSI та CVSSI займає ще більше часу (приблизно на порядок). Таким чином, навіть при паралельних обчисленнях елементарних метрик, обчислення метрики на основі нейромережі



потребує значно більше часу, ніж для таких хороших елементарних метрик, як MDSI, GMSD або HaarPSI. Отже, подальші напрямки досліджень повинні бути спрямовані на поєднання швидких елементарних метрик, забезпечуючи хороший баланс між монотонністю прогнозування MOS (а також точністю прогнозування) та обчислювальною ефективністю.

Таким чином, нами розглянуто завдання оцінки візуальної якості зображень дистанційного зондування з різними типами спотворень. Показано, що не існує загально визнаних метрик візуальної якості, і, отже, бажаною є їх розробка. Зазначено проблеми, з якими можна зіткнутися, та запропоновано їх вирішення. Для цієї мети використана вже існуюча база даних спотворених кольорових зображень, для яких можна вибрати зображення із типовими спотвореннями, які часто спостерігаються при дистанційному зондуванні. Її використання дозволило визначити існуючі метрики візуальної якості, які добре працюють для видів спотворень, що нас цікавлять. Найкращі з таких показників забезпечують SROCC з MOS близько 0,93, що вважається дуже хорошим результатом. Більше того, база даних TID2013 дозволяє розробляти метрики візуальної якості на основі використання елементарних метрик як вхідних даних. Було вивчено декілька конфігурацій нейромереж та методи попередньої обробки вхідних даних. Було показано, що навіть прості нейромережі без попередньої обробки вхідних даних (лінеаризації), що мають два приховані шари, здатні надати значення SROCC близько 0,97. Значення PCC мають той же порядок, що означає, що зв'язок між вихідним сигналом нейромережі та значеннями MOS є практично лінійним. Потім деякі елементарні та розроблені метрики були перевірені для триканальних зображень дистанційного зондування з двома типами спотворень. Перевірка продемонструвала монотонність поведінки запропонованої метрики. Крім того, було показано, що розроблена метрика забезпечує точну оцінку MOS, що дозволяє класифікувати зображення дистанційного зондування відповідно до їх якості.

Надалі планується провести аналіз способів прискорення метрик на основі нейромереж за допомогою обмеження набору можливих метрик на вході, враховуючи обчислювальну ефективність вхідних метрик. Ми також сподіваємось, що будуть проведені деякі експерименти щодо сприйняття за допомогою фахівців з аналізу зображень ДЗ для оцінки якості зображення.

Інший напрямок наших подальших досліджень буде пов'язаний з оптимізацією методів перетворення кольорів, що використовуються для окремих



елементарних метрик. У деяких випадках відповідне застосування елементарних метрик для багатоканальних зображень може також вимагати змін типів даних та динамічних діапазонів. Результати окремих елементарних метрик, отримані з використанням різних колірних просторів та методів перетворення, можуть призвести до подальшого підвищення ефективності комбінованих метрик.

### **4.3. Оцінка якості тривимірних друків поверхонь з використанням комбінованих метрик**

Тривимірний (3D) друк є однією з ключових технологій, що революціонізує невелике серійне виробництво в епоху промисловості 4.0. Використання 3D-принтерів дозволяє не лише створити деякі оригінальні 3D-об'єкти для розважальних цілей, але й запустити індивідуальне виготовлення деяких унікальних деталей машин та інших пристроїв, що використовуються для заміни деяких пошкоджених старих елементів. Є й інші галузі застосування технології адитивного виробництва із використанням пластикових ниток, як правило, на основі поліактової кислоти (PLA) або акрилонітрилбутадієнстиролу (ABS). Вони можуть бути пов'язані з біомедичною технікою (наприклад, індивідуальним протезуванням), аерокосмічною та автомобільною промисловістю, цивільним будівництвом та архітектурою, зворотнім проектуванням у промисловості або навіть охороною культурної спадщини. Інтеграція з 3D-сканерами дозволяє досить легко створювати власні 3D-зображення CAD-моделі та копії різних елементів. Незважаючи на зростаючу популярність 3D-друку для домашнього використання слід враховувати деякі важливі обмеження, такі як викиди частинок [89, 90], особливо з використанням ABS-ниток.

Тим не менше, на процес 3D-друку можуть впливати різні фактори, що впливають на остаточний результат та якість виготовлених предметів, особливо для недорогих пристроїв, призначених для домашнього використання. Такі принтери належать до найпопулярнішої групи, заснованої на моделюванні плавленого осадження (FDM), де нагріту нитку наносять знизу вгору шар за шаром рухомим екструдером, що утворює 3D друкований об'єкт. Одним з таких джерел поверхневих спотворень може бути неправильна температура плавлення, що залежить від типу нитки, а також деяких змін температури навколишнього середовища. Деякі інші проблеми можуть бути пов'язані з низькою якістю



елементів, що використовуються для конструкції принтера, та низькою якістю ниток. Деякі спотворення можуть також бути спричинені неправильною конфігурацією степера, а також зміною швидкості подачі нитки. Обидва ці типи спотворень можуть з'явитись під час підготовки випробувальних зразків, які були використані для розробки бази даних, що використовується для перевірки запропонованих методів [91, 92].

Зростання популярності та доступності тривимірних принтерів, а також відносно висока роздільна здатність дешевих камер дозволяють інтегрувати деякі алгоритми комп'ютерного зору, які можуть бути корисними для моніторингу процесу друку в реальному часі (тобто, без помітних затримок відносно швидкості виготовлення) та стану пристрою [93, 94]. Тим не менше, більшість відомих рішень обмежується спостереженням за станом пристрою і застосовується переважно для діагностики несправностей [95–97]. Деякі інші рішення передбачають використання довідкових даних, що представляють деякі особливості або дескриптори, такі як підписи процесу [98, 99]. Деякі із запропонованих систем використовують оптичну когерентну томографію (ОСТ) [100], термографічні вимірювання [101] або терагерцову технологію [102]. Важливі аспекти виявлення деяких проблем якості можуть бути пов'язані з кібербезпекою виробничих систем [103, 104], особливо враховуючи, що в деяких випадках деякі дефекти в зразках, надруковані у форматі 3D, можуть залишитися невизначеними, наприклад, за допомогою ультразвукового контролю [105].

Оскільки використання складних апаратних рішень для моніторингу недорогих 3D-принтерів є проблемним через значне збільшення загальних витрат системи, то найбільш розумним рішенням для любительського використання представляється аналіз зображень, отриманих доступними камерами. Одна з таких системи для моніторингу виробничого процесу на основі порівняння надрукованої поверхні з комп'ютерною 3D-моделлю була запропонована Хольцмондом та Лі [105]. Ще одна спроба виявлення аномалій та класифікації методу лазерного порошкового плавлення (LPBF) на основі некерованого алгоритму машинного навчання була зроблена у роботі [106]. Але такий автоматизований аналіз вимагає підготовки алгоритму, обмежуючи його практичну придатність у недорогих пристроях. Подібна проблема також характерна для використання нейронних мереж, використаних, наприклад, для моніторингу процесу 3D-струменевого друку електронних виробів [107].

Цікаве застосування машинного зору для виявлення дефектів у 3D-об'єктах



було нещодавно запропоновано Штраубом [108]. Це рішення, засноване на п'яти камерах та блоках Raspberry Pi, призначено для зменшення або усунення необхідності тестування друкованих об'єктів через можливість автоматичного виправлення незначних дефектів, помічених в процесі друку, а також виявлення проблем “сухого друку”, спричинених відсутністю нитки матеріалу. Але запропонована система може виявитися дуже чутливою до умов навколишнього середовища, а також навіть до незначних рухів камери. Застосування видимого зондування для виявлення мікрodefektів у 3D-друкованих об'єктах, що використовують програми, що мають найважливіше значення для безпеки, представлені в роботі [109].

Оскільки практична придатність суб'єктивних методів оцінювання якості друку сильно обмежена часом, необхідним для того, щоб набирати бали для багатьох зображень, як це робиться для фільтрації зображень та їх стиснення з втратами, то єдиною можливістю є використання об'єктивних метрик, які можуть обчислюватися автоматично без участі людей-спостерігачів. Такі метрики (показники) можна розділити на три сімейства: метрики без еталону, які також називаються "сліпими" метриками, які не вимагають знання вихідного зображення, метрики з використанням часткової довідкової інформації та найпопулярніші метрики з еталоном (FR IQA), що реалізують порівняння спотворених зображень з “первинними” еталонними зображеннями без будь-яких спотворень. Такі еталонні зображення ідеальної якості зазвичай включені до баз, що є корисними головним чином при розробці деяких нових метрик з еталоном. Незважаючи на потенційно широкі сфери застосування метрик без еталону, їх універсальність все ще значно нижча порівняно з метриками з еталоном, що спричиняє більшу популярність останніх підходів.

Більшість загальних методів оцінки якості з еталоном засновані на припущенні, що обидва зображення представляють одну і ту ж сцену, але одне із зображень пошкоджене одним або кількома типами спотворень (див. підрозділ 4.1). При цьому під час порівняння зображень припускається, що зображення не мають взаємного зсуву чи обертання.

Аналіз властивостей елементарних метрик можна знайти у багатьох роботах [40-42, 110] та у підрозділі 4.1. У попередніх підрозділах показано, як їх можна об'єднати, що призводить до багатообіцяючих результатів. Тим не менше, пряме застосування такого підходу для оцінки якості тривимірних друкованих об'єктів відбитків може бути складним через час, необхідний для обчислення



кількох метрик, а також внаслідок необхідності навчання нейронної мережі. Враховуючи різноманітність запропонованих підходів на основі подібності оцінки якості зображення, нижче перевірена корисність метрик для автоматичної оцінки якості поверхонь з 3D-друком, а також можливості їх комбінування із потенційним застосуванням в системах реального часу.

Перевірка корисності різних об'єктивних метрик для оцінки якості 3D-друку поверхонь вимагає розробки спеціальної бази даних, що містить зображення різних спотворених і високоякісно виготовлених поверхонь разом із суб'єктивними значеннями MOS. З цією метою були виготовлені плоскі зразки з 9 типів термопластичних ниток різних кольорів із використанням трьох доступних пристроїв FDM, а саме RepRap Ormerod 3, Prusa i3 та XYZprinting da Vinci 1.0 Pro 3-в-1. Вибір ABS ниток був мотивований їх хорошими механічними властивостями і легкістю, а також вищою стійкістю до стирання в порівнянні з полімерами PLA. Тим не менше, відносно високі температури плавлення - понад 220°C - разом із потенційно токсичними випарами трохи обмежили розвиток цього набору даних.

Хоча база даних містить 107 фотографій (разом із картами глибин, отриманими 3D-сканером, який не використовується в даній роботі) плоских поверхонь, метрики, що досліджуються в цій роботі, можуть бути успішно застосовані незалежно від плоскості поверхні; отже, в цьому сенсі запропонований підхід може розглядатися як універсальний і може бути додатково перевірений також для деяких інших поверхонь.

Зображення були отримані в контрольованому освітленні (використовувалось розподілене освітлення від трьох лампи для запобігання сильних відблисків) за допомогою камери Sony DSC-HX100V із часом витримки 1/125 с, фокусна відстань дорівнювала 5 мм, використовувався автоматичний баланс білого без спалаху, забезпечуючи фіксовану відстань до поверхні. Розмір зображень становить 1600 × 1600 пікселів, що еквівалентно фізичному розміру зразків 35 мм × 35 мм. Їх товщина становить близько 4 мм, а висота шарів варіюється від 0,3 до 0,35 мм в залежності від принтера та розміру сопла. Карти глибини були отримані за допомогою 3D-сканера ATOS компанії GOM.

Незалежно від впливу деяких незалежних факторів, таких як якість нитки та матеріалу, наявність спотворень у деяких зразках було зумовлено зміною температури, параметрами конфігурації крокових двигунів, швидкістю подачі нитки. Деякі зразки містили тріщини, а також результати недо- та переповнення.





Усі фотографії були незалежно оцінені 92 спостерігачами, використовуючи п'ятибальну шкалу якості від 1 (дуже погано) до 5 (дуже добре). Отримані результати були усереднені для отримання значень MOS, що додатково перевіряється шляхом порівняння з попередніми думками від експертів (такі думки (оцінки) використовувались для класифікації в деяких попередніх роботах). Один зразок зображення разом зі значеннями MOS представлені на рисунку 12, де значення MOS, що наближаються до 5, відповідають ідеальній якості.

Ідея поєднання різних ознак частково використовується в багатьох обговорюваних метриках, починаючи з SSIM, що фактично являє собою комбінацію трьох компонентів, що представляють яскравість, контрастність і структуру. Тим не менше, така комбінація також може бути застосована для окремих (елементарних) метрик, що зазвичай призводить до значного збільшення кореляції з суб'єктивними оцінками якості. Найкращі результати можуть бути отримані для комбінації метрик, що використовують різні типи ознак, доповнюючи одна одну. Різні приклади об'єднання наведені в роботах [111-117].

Втім, розглядаючи застосування метрик для оцінки якості 3D-друку поверхонь під час виробничого процесу, доцільним здається використання обмеженої кількості метрик у поєднанні з використанням вагових коефіцієнтів згідно за загальною формулою:

$$Q_{combined} = \prod_{n=1}^N Metric_n^{weight_n}, \quad (5)$$

де  $N$  - кількість зважених елементарних метрик.

Оскільки метрики, починаючи з UIQI та SSIM, належать до групи методів оцінювання якості з еталоном, їх безпосереднє застосування для розглянутого застосування вимагало б знання еталонного зображення. У практичному застосуванні це, як правило, неможливе і навіть використання візуалізованої моделі тривимірно надрукованого об'єкта вимагатиме точної реєстрації зображень. Додаткова проблема може бути пов'язана з різними кольорами та яскравістю отриманих зображень та зображень з 3D-моделей. Отже, пропонується ідея обчислення середньої взаємної подібності зображень поверхонь з 3D друком, що передбачає бічне розташування камер. Спочатку такий варіант досліджувався у статтях [118, 119] виключно для цілей класифікації. Таке кріплення камери збоку дозволяє отримати зображення із



видимими окремими шарами нитки, як показано на рисунку 12.



- а) високоякісний зразок жовто-рожевого кольору (MOS = 4,7253);  
 б) зразок високої якості рожевого кольору (MOS = 4,6923);  
 в) високоякісний зразок коричневого кольору (MOS = 4,1333);  
 г) зразок помірно високої якості червоного кольору (MOS = 2,4130);  
 д) зразок помірно низької якості жовтого кольору (MOS = 1,4130);  
 е) зразок низької якості темно-зеленого кольору (MOS = 1,1868);  
 є) низькоякісний чорний зразок (MOS = 1,0978);  
 ж) низькоякісний зразок рожевого кольору (MOS = 1,0110);  
 з) низькоякісний зразок блакитного кольору (MOS = 1,0000).

**Рисунок 12 – Зразки зображень з розробленої бази даних із середніми суб'єктивними показниками якості із сортуванням від найвищого до найнижчого**

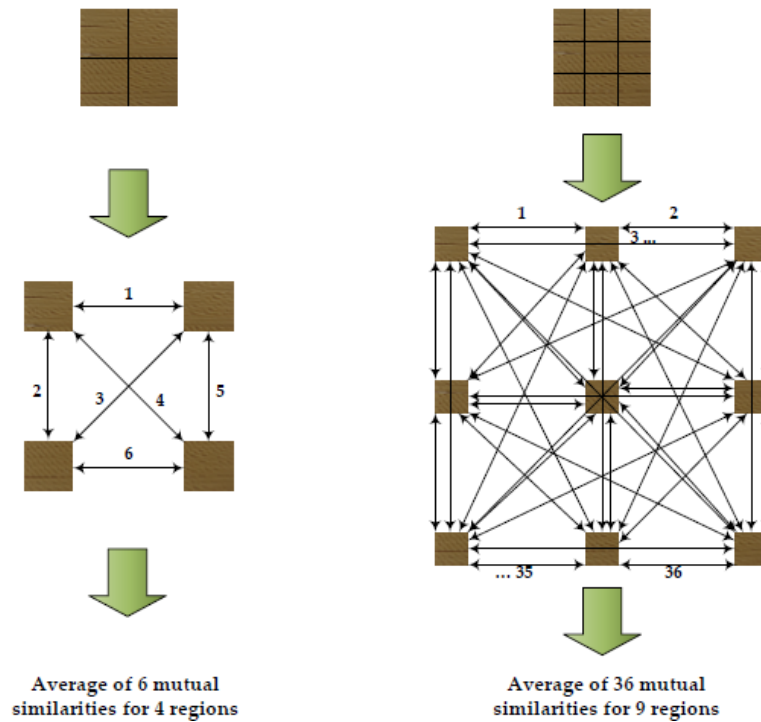


Ідея обчислення взаємної подібності вимагає поділу зображення, що представляє виготовлений зразок, на регіони та застосування формули оцінювання якості для кожної пари таких отриманих блоків. У проведених експериментах все зображення було розділено на 4, 9 та 16 квадратних блоків. Прогнозувалося, що найкращі результати будуть отримані для 16 блоків (4×4 областей), що вимагає 120 взаємних порівнянь. Оскільки розподіл поверхні на області фіксований, ми можемо очікувати такі ж або дуже схожі орієнтації ділянок у кожному регіоні. Отже, завдяки використанню взаємної подібності на основі порівняння регіонів, потенційний вплив позиції та орієнтації регіону було значно зменшено. Завдяки застосуванню обробки у ковзному вікні для метрик якості на основі SSIM, структурні зміни, внесені такими порівняно невеликими обертами або зсувами впливають на кінцеві результати розрахунку окремих (елементарних) метрик не так сильно, як наявність фізичних спотворень поверхні 3D-друку.

Крім того, може бути застосовано фазовий зсув на кілька пікселів для коригування фази порівняних моделей. Вплив кольору на окремі зразки зменшився за рахунок використання перетворення кольорових даних у зображення в градаціях сірого відповідно до рекомендації МСЕ ITU-R BT.601-7 (з використанням функції `rgb2gray` в MATLAB R). Ілюстрація ідеї взаємної подібності представлена на рисунку 13, де представлені два приклади поділу на 4 та 9 регіони з необхідними 6 та 36 обчисленнями взаємної подібності.

Отже, середні значення взаємної подібності, отримані для елементарних метрик, можуть розглядатися як відповідні еквіваленти метрик з еталоном, що застосовуються для оцінки якості 3D поверхонь. Оскільки ми не використовуємо “первинні” контрольні зображення, то немає можливості застосовувати методи оцінки якості з еталоном безпосередньо. Тож зображення були розділені на частини і проведено порівняння цих регіонів, щоб усунути відсутність еталонних зображень. Отже, формально можна класифікувати наш підхід як підхід до оцінки без еталону („сліпий”), хоча насправді всі метрики (показники), що використовуються «в середині» належать до групи методів з еталоном.

Такі отримані показники якості використовувались як вхідні показники для комбінованої метрики (5). Оптимізовані ваги були отримані шляхом максимізації коефіцієнту лінійної кореляції Пірсона (PLCC) між об'єктивною та суб'єктивною оцінкою (виражається як значення MOS у розробленій базі даних). Оскільки було припущено, що нелінійна комбінація елементарних метрик

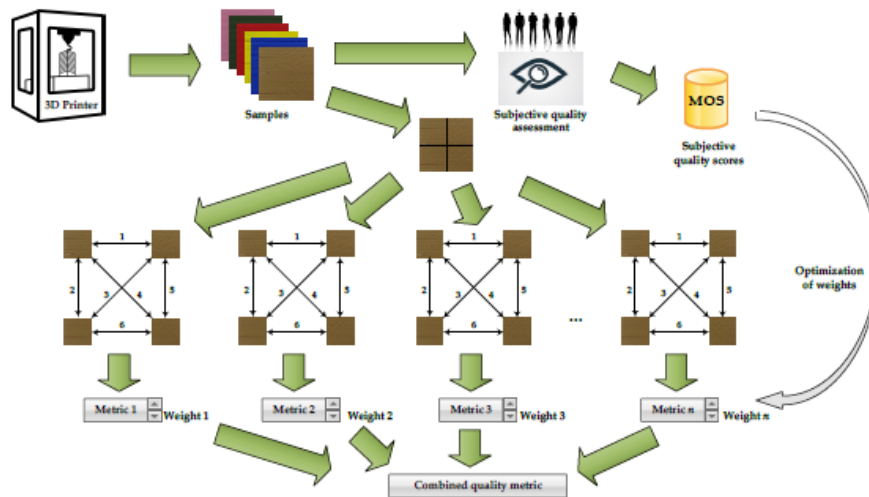


**Рисунок 13 – Ілюстрація ідеї розрахунку взаємної подібності з поділом на 4 та 9 регіонів з необхідними відповідними 6 та 36 обчисленнями метрики з еталоном**

повинна компенсувати потенційно нелінійність сприйняття спотворень, то не застосовувались жодні функції для фітінгу. Крім того, було розраховано дві рангові кореляції, а саме ранговий коефіцієнт кореляції Спірмена (SROCC), позначений як  $\rho$ , та ранговий коефіцієнт кореляції Кендалла (KROCC), позначений як  $\tau$ .

Перші експерименти були проведені, передбачаючи використання окремих елементарних метрик для всіх 107 зображень, перетворених у відтінки сірого та розділених на 4, 9 та 16 областей. Для кожного з них були розраховані три коефіцієнти кореляції для виявлення потенційно найбільш корисних показників для подальшого використання під час оптимізації комбінованих метрик.

Отримані результати представлені в таблиці 12. Як можна помітити, в більшості випадків для окремих метрик були отримані подібні значення PLCC і SROCC, отже, в подальших експериментах в якості критерію оптимізації використано лише кореляцію Пірсона. Для багатьох метрик вибір кількості регіонів істотно не змінює значення кореляції, отже, подальші експерименти були проведені за умови поділу на 16 регіонів.



**Рисунок 14 – Ілюстрація експериментальної процедури**

Найкращі результати PLCC, отримані для різних комбінацій двох метрик представлені в таблиці 13 (для економії місця наведені лише дані із значеннями вище 0,67). Як видно, у більшості комбінацій використовувалась метрика FSIM, отже, цю метрику нами було використано як "фіксовану" для комбінацій з трьох метрик (змінювались лише дві інші під час оптимізації ваг, вагу для FSIM також було оптимізовано). Найкращі результати PLCC, досягнуті для таких комбінацій трьох метрик, представлені в таблиці 14.

**Таблиця 12 – Коефіцієнти кореляції, отримані для 107 зображень в базі, з використанням запропонованої взаємної подібності та елементарних метрик**

Metric	Division into 16 Regions			Division into 9 Regions			Division into 4 Regions		
	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC	PLCC	SROCC	KROCC
FSIM[34]	0,6756	0,6865	0,5195	0,6820	0,6845	0,5185	0,6780	0,6826	0,5114
CW-SSIM[32]	0,5929	0,5823	0,4028	0,6323	0,6098	0,4232	0,5807	0,5633	0,3981
AD-GSIM[40]	0,4081	0,3515	0,2453	0,4020	0,3470	0,2414	0,3873	0,3324	0,2354
DSS [38]	0,4066	0,3523	0,2411	0,3842	0,3220	0,2210	0,3921	0,3176	0,2142
AD-SSIM[40]	0,4017	0,3574	0,2562	0,3834	0,3209	0,2270	0,3492	0,2932	0,2065
SSIM [29]	0,3996	0,4012	0,2746	0,3905	0,4039	0,2661	0,3048	0,3017	0,1938
CSSIM4[41]	0,3596	0,3296	0,2354	0,3329	0,2818	0,1977	0,3233	0,2851	0,1991
IW-SSIM[33]	0,3549	0,3669	0,2619	0,3230	0,2997	0,2044	0,3169	0,2473	0,1627
SR-SIM[36]	0,3173	0,2441	0,1588	0,3174	0,2497	0,1652	0,3878	0,3160	0,2150
MCSID [39]	0,3106	0,2958	0,2164	0,3008	0,2889	0,2090	0,2952	0,2825	0,2051
QILV [31]	0,3092	0,1330	0,0868	0,3478	0,2662	0,1832	0,4316	0,3555	0,2534
CVSSI [42]	0,2558	0,2083	0,1370	0,2097	0,1492	0,0935	0,1667	0,1068	0,0593
ESSIM [37]	0,1865	0,2340	0,1631	0,1754	0,2354	0,1648	0,3160	0,2868	0,2026
CSSIM [41]	0,1523	0,1078	0,0724	0,1251	0,0755	0,0519	0,1293	0,0862	0,0632
SSIM4 [41]	0,1283	0,0852	0,0565	0,1031	0,0447	0,0304	0,1085	0,0673	0,0462
GSM [35]	0,1103	0,1689	0,1182	0,0991	0,1631	0,1133	0,2102	0,2253	0,1585
RVSIM [43]	0,0267	0,0198	0,0219	0,0546	0,0433	0,0395	0,0114	0,0247	0,0304



**Таблиця 13 – Значення PLCC, отримані для 107 зображень, з використанням 25 найкращих комбінацій двох елементарних метрик (поділ на 16 регіонів)**

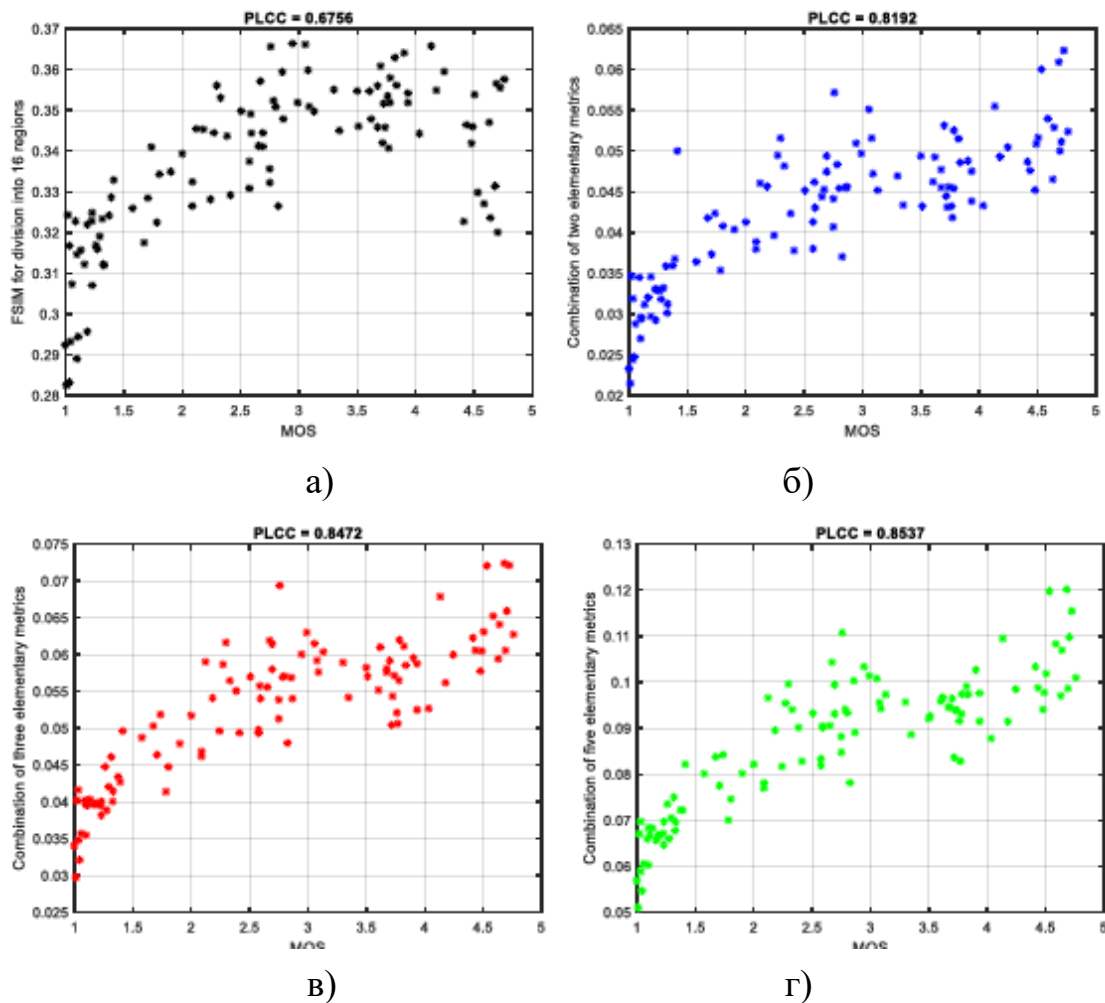
Metric	PLCC	Metric	PLCC	Metric	PLCC	Metric	PLCC
FSIM+MCSD	0,8192	FSIM+SR-SIM	0,7862	FSIM+SSIM4	0,7581	FSIM+RVSIM	0,6875
FSIM+DSS	0,8029	FSIM+AS-SSIM	0,7861	FSIM+ESSIM	0,7377	CSSIM+CSSIM4	0,6852
FSIM+CVSSI	0,8008	FSIM+CW-SSIM	0,7766	FSIM+GSM	0,7312	FSIM+SSIM	0,6809
- FSIM+IW-SSIM	0,7981	FSIM+CSSIM4	0,7719	SSIM+AD-GSIM	0,7081	FSIM+RVSIM	0,6762
FSIM+AD-GSIM	0,7921	FSIM+CSSIM	0,7593	FSIM+QILV	0,6952	SSIM+CW-SSIM	0,6731

**Таблиця 14 – Значення PLCC, отримані для 107 зображень, для 24 найкращих комбінацій з трьох елементарних метрик (поділ на 16 регіонів)**

Metric	PLCC	Metric	PLCC	Metric	PLCC
FSIM+CW-SSIM+MCSD	0,8472	FSIM+MCSD+GSM	0,8270	FSIM+MCSD+AD-SSIM	0,8221
FSIM+CW-SSIM+DSS	0,8379	FSIM+ESSIM+MCSD	0,8256	FSIM+MCSD+CSSIM	0,8221
FSIM++CW-SSIM+IW-SSIM	0,8356	FSIM+CW-SSIM+CSSIM4	0,8250	FSIM+CVSSI+GSM	0,8219
FSIM+CW-SSIM+CVSSI	0,8348	FSIM+CW-SSIM+SR-SIM	0,8246	FSIM+MCSD+RVSIM	0,8217
FSIM++AD-SSIM+SR-SIM	0,8341	FSIM+CW-SSIM+AD-SSIM	0,8239	FSIM+DSS+AD-SSIM	0,8213
FSIM+CW-SSIM+AD-GSIM	0,8301	FSIM+MCSD+SR-SIM	0,8238	FSIM+ESSIM+AD-GSIM	0,8213
FSIM+DSS+ESSIM	0,8284	FSIM+MCSD+SSIM4	0,8225	FSIM+MCDS+QILV	0,8201
FSIM+DSS+GSM	0,8274	FSIM+MCSD+SSIM	0,8224	FSIM+AD-GSIM+GSM	0,8199

Результати, представлені в таблиці 14, свідчать, що найкращі результати можна отримати для комбінації FSIM та CW-SSIM з третьою метрикою (бажано MCSD), однак хороші результати також можуть бути досягнуто за допомогою FSIM та MCSD як двох основних метрик, що використовуються у комбінаціях. Враховуючи ці результати, були розглянуто ще більша кількість елементарних метрик. Це призвело до ще кращої кореляції комбінованих метрик зі значеннями MOS, але покращення вже малі у порівнянні з комбінованими метриками на основі трьох елементарних метрик. Найкращі значення PLCC не завжди еквівалентні найвищим значенням SROCC та KROCC. Втім, такі відмінності можна вважати незначними.

Додаткова ілюстрація отриманого збільшення кореляції з суб'єктивними балами представлена на скатерограмах на рисунку 15.



- а) для FSIM; б) для "найкращої" комбінації з двох метрик;
- в) для "найкращої" комбінації з трьох метрик;
- г) для "найкращої" комбінації з п'яти метрик.

**Рисунок 15 – Скатерограми, що ілюструють лінійність співвідношень між досліджуваними об'єктивними метриками (поділ на 16 регіонів) та значенням MOS**

Хоча перевірка кореляцій для комбінацій двох та трьох метрик розрахована, припускаючи поділ на 4 або 9 блоків, дослідження підтвердили обґрунтованість вибору 16 регіонів, що призводить до кращої ефективності. Поєднання чотирьох та п'яти метрик для меншої кількості блоків дає можливість отримати трохи кращі результати, враховуючи монотонність відповідно до обох коефіцієнтів рангової кореляції. Використання меншої кількості регіонів зменшує час обчислення, але обчислення більшої кількості метрик збільшує його, тож можна пошукати розумний компроміс.

Представлені результати підтверджують корисність комбінованих метрик



для оцінки якості 3D друків поверхонь, що призводить до високої кореляції з суб'єктивними думками щодо естетики поверхні. Запропоновано підхід до автоматичного оцінювання якості об'єктів поверхонь 3D-друка з використанням лише метрик з еталоном, які покладаються на взаємну подібність. Це дозволяє порівняно легко розрахувати значення елементарних метрик, необхідних для комбінації. Крім того, вони можуть обчислюватись паралельно, що дозволяє додатково прискорити обробку.

## Висновки

Під час виконання досліджень головну увагу приділено розробці комбінованих метрик якості з еталоном, їх оптимізації та аналізу можливостей нових застосувань – при аналізі якості багатоканальних даних дистанційного зондування та якості поверхонь під час тривимірного друку. На основі проведеного аналізу та розробок можна зробити наступні висновки:

а) існує можливість суттєво підвищити кореляцію метрик візуальної якості та MOS за рахунок як мінімум трьох підходів до розробки комбінованих метрик: на основі стійких оцінок й попередньої лінеаризації, на основі навчання нейромереж та використовуючи оптимізації ваг для формули (5);

б) нейромережі дають найкращі результати, але вимагають навчання (офлайн) та використання елементарних метрик, частина з яких може мати низьку швидкість, що гальмує розрахування комбінованої метрики; в цьому сенсі є напрямки подальшої роботи, що може призвести до прискорення без втрати ефективності обробки даних нейронною мережею;

в) розроблені підходи до синтезу комбінованих метрик є досить універсальними; ми прогнозуємо, що є багато інших застосувань, де такі метрики будуть ефективними;

г) значення кореляції або перевищують 0,9 (сягаючи 0,97 для нейромереж), або наближуються до 0,9, що свідчить про високу адекватність розроблених метрик.

Отримані результати більш детально відображено у публікаціях авторів розділу [19-21].