

П.А. КОВАЛЬ

Національний університет «Львівська політехніка», Львів, Україна,
e-mail: pavlo.koval.mknus.2024@lpnu.ua.

І.О. ПРОЦЬКО

Національний університет «Львівська політехніка», Львів, Україна,
e-mail: ihor.o.protsko@lpnu.ua.

ДОСЛІДЖЕННЯ ЗАСТОСУВАННЯ БАЗОВОЇ РЕКОНСТРУКТИВНОЇ ТА УМОВНОЇ GAN-МОДЕЛЕЙ НЕЙРОМЕРЕЖ ДЛЯ РЕАЛІЗАЦІЇ ПЕРЕТВОРЕНЬ ЗОБРАЖЕНЬ LWIR-RGB

Анотація. У роботі описано дослідження та порівняння перетворення зображень довгохвильового інфрачервоного діапазону у видиме представлення (LWIR-RGB) з використанням реконструктивної базової моделі U-Net та умовної GAN-моделі pix2pix з локальним дискримінатором. Програмну платформу реалізовано у вигляді відтворюваного контейнера перетворення зображень на основі архітектур нейронних мереж U-Net та pix2pix, що застосовуються у задачах навчання з учителем типу «зображення–зображення». Наведено результати експериментального дослідження перетворень зображень довгохвильового інфрачервоного діапазону у видиме представлення U-Net та pix2pix моделей, навчених на парних даних із набору даних KAIST. Експерименти ґрунтуються на оцінюванні та порівнянні комбінованих функцій втрат з ваговими коефіцієнтами для індексу структурної подібності та середньої абсолютної різниці між передбаченим і еталонним зображеннями. Розроблені комплексні функції втрат поєднують розуміння глобального контексту зображення зі збереженням локальних ознак. Наведено графіки навчання двох (U-Net та pix2pix) моделей нейромереж, які дають якісне уявлення про динаміку навчання та його збіжність. Відповідно до вибраного оцінювання добре навчена U-Net-мережа дає хорошу структурну реконструкцію. Змагальна складова моделі pix2pix з використанням локального дискримінатора PatchGAN забезпечує реалістичне створення зображень на різних рівнях деталізації, покращуючи сприйнятву реалістичність без втрати структурної якості перетвореного зображення. Визначено подальші напрями покращення достовірності перетворення одноканальних кадрів теплових зображень у видимі триканальні RGB-представлення на основі досліджених моделей.

Ключові слова: перетворення «зображення–зображення», теплові зображення, глибинне навчання, тензор зображення.

DOI 10.34229/KSA2522-9664.26.3.5

ВСТУП

Протягом останнього десятиліття перетворення зображень між різними модальностями стало однією з активно досліджуваних задач глибинного навчання. Значна кількість робіт присвячена розробленню архітектур, функцій втрат і стратегій навчання для задач типу «зображення–зображення», зокрема у випадках, коли вхідні та вихідні дані належать до різних фізичних доменів. Одним із типових прикладів такої міжмодальної задачі є перетворення зображень довгохвильового інфрачервоного діапазону (LWIR) у видимі RGB-представлення.

Теплові та видимі зображення ґрунтуються на принципово різних фізичних явищах. Зображення LWIR відображають розподіл теплового випромінювання та температурні характеристики об'єктів, тоді як зображення у видимому спектрі формуються за рахунок відбитого світла та містять інформацію про колір та дрібну текстуру поверхні. Через відсутність кольорової та текстурної інформації у теплових даних задача перетворення LWIR-RGB є за своєю природою недовизначеною: одному тепловому зображенню може відповідати кілька можливих видимих інтерпретацій. Це створює фундаментальний компроміс між прагненням до візуальної правдоподібності та необхідністю збереження семантичної і структурної коректності результату [1].

© П.А. Коваль, І.О. Процько, 2026