

УДК 004.932

В. Попсуй; В. Бревус, к. т. н.

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

СИСТЕМА КОНТРОЛЮ ЯКОСТІ СКАНІВ У MAYAN EDMS З ВИКОРИСТАННЯМ OPENCV ТА МАШИННОГО НАВЧАННЯ

UDC 004.932

V. Popsui; V. Brevus, PhD.

QUALITY CONTROL SYSTEM FOR SCANS IN MAYAN EDMS USING OPENCV AND MACHINE LEARNING

Цифровізація документообігу та зростання обсягів архівних фондів посилюють потребу в автоматизації контролю якості сканованих документів. Сучасні дослідження доводять ефективність методів комп'ютерного зору у виявленні дефектів цифрових зображень – розмиття, перекосів, низької контрастності та артефактів [1–3]. Додатково, розвиток легковагових моделей глибинного навчання та підходів до класифікації візуальних даних у контексті Industry 4.0 [4–7] відкриває можливість інтеграції автономних систем контролю у корпоративні EDMS-платформи.

Об'єктом дослідження є процес завантаження та індексування документів у системі Mayan EDMS. Виявлення технічних дефектів реалізовано за допомогою методів OpenCV: Laplacian-variance для оцінки різкості, Hough-трансформацій для визначення перекосу, гістограмного аналізу для виявлення проблем освітлення. Для класифікації складних дефектів – зім'яття сторінки, часткової втрати фрагментів, сторонніх об'єктів.

Розроблена система інтегрується як додатковий модуль у Mayan EDMS та аналізує якість сканів на ключових етапах: під час завантаження документа, після обробки OCR та перед індексацією. Автоматична логіка забезпечує три напрямки адаптивної взаємодії:

- **попереджувальний контроль** – блокування низькоякісних сканів та рекомендації повторного сканування;
- **інтелектуальне оцінювання** – формування оцінки якості за комплексом метрик та виявлення специфічних дефектів;
- **автоматизована оптимізація** – корекція контрасту, вирівнювання перспективи та нормалізація яскравості.

Експериментальні результати підтвердили, що поєднання класичних алгоритмів OpenCV із легковаговими ML-моделями забезпечує високу точність виявлення дефектів при мінімальному навантаженні на систему. Це підвищує надійність зберігання електронних документів та зменшує людський фактор у процесах цифрового архівування.

Література

1. Bradski G. The OpenCV Library // Dr. Dobb's Journal. – 2000.
2. Shafait F., Keysers D., Breuel T. Performance Evaluation and Benchmarking of Six Page Segmentation
3. Chen X., Zhang L., Doermann D. Document Image Quality Assessment: A Brief Survey // 2015 13th International Conference on Document Analysis and Recognition (ICDAR). – 2015.
4. Zhou S., He L., Wang W. Deep Learning-Based Document Image Quality Assessment // Pattern Recognition. – 2022.
5. Smith R. An Overview of the Tesseract OCR Engine // ICDAR. – 2007.
6. Mayan EDMS Documentation. Official Technical Documentation.