

КРИВЕНЧУК Ю. П.

Національний університет "Львівська політехніка"

<https://orcid.org/0000-0002-2504-5833>e-mail: [Yurii.P.Kryvenchuk@lpnu.ua](mailto:Yurii.P.Kryvenchuk@lpnu.ua)

ЯКИМІВ Р. Ю.

Національний університет "Львівська політехніка"

e-mail: [roman.yakymiv.knm.2018@lpnu.ua](mailto:roman.yakymiv.knm.2018@lpnu.ua)

## СИСТЕМА АНАЛІЗУ ОБЛИЧЧЯ ЛЮДИНИ ДЛЯ АПРОКСИМАЦІЇ ВІКУ

В роботі наведено результати дослідження теми апроксимації віку людини за її обличчям та розроблення системи для вирішення такої задачі, з метою покращення якості та збільшення швидкодії такого типу систем, їх поширення та впровадження у відповідні сфери в Україні. Виділено та описано такі основні етапи: розпізнавання обличчя, визначення ознак, апроксимація віку. Після проведення аналізу результатів було виявлено, що створення системи апроксимації віку людини за зображенням її обличчя є актуальним та доцільним завданням на сьогодні, а найбільш ефективним інструментом для цього є використання методів глибокого навчання та згорткових нейронних мереж.

Ключові слова: апроксимація віку, згорткові нейронні мережі, обличчя.

Yurii KRYVENCHUK, Roman YAKYMIV  
Lviv Polytechnic National University

## CREATION OF AGE ESTIMATION SYSTEM

We are surrounded by different transmitting and processing information devices on a daily basis. Analysis of certain selected characteristics of each person allows humanity to achieve previously unthinkable technological development. In particular, the human face. We can predict a person's age, determine his mood or identify someone in the crowd using modern algorithms. As a "window to the soul", the human face provides important information related to its individual traits. With the help of this information, a person can determine such characteristics as ethnic origin, gender, age, and emotions, on the basis of which he is able to take appropriate action, which plays a significant role in non-verbal communication between people. Age estimation is useful in applications where you do not need to specifically identify a person, but we want to know (approximate) their age. With the help of the age approximation system, it is possible to ensure partial digitalization of security control and monitoring systems. With the help of a monitoring camera, the age estimation system can identify minors and prevent them from entering places where they are not allowed; not to allow minors to buy tobacco products or alcohol from vending machines; to warn the elderly person of possible danger. Age assessment software can also be used in healthcare systems. The authors personally are interested in approximating human age based on images to personalize human-machine interaction. The paper has the description and the results of the creation of an age estimation system, the basis of which is a convolutional neural network. The user can connect with the system through a simple interface which allows you to upload the image for analysis or turn on the camera to get age estimates in real-time. The practical value of this work is the developed high-quality system which is ready for use and implementation in the relevant fields.

Keywords: age approximation, convolutional neural networks, face.

### Постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями

Вік є фундаментальною характеристикою людини, визначивши та проаналізувавши яку ми отримуємо змогу швидко дізнатись важливу інформацію про особу чи групу людей. Точно визначити вік людини за виглядом неможливо, тому буде апроксимовано її вікову категорію. Інформацію, яку надасть автоматизована система апроксимації віку можна використовувати для автоматизованого збору та аналізу даних, у безпекових та контрольних цілях, для персоналізації людино-машинної взаємодії.

Дану розробку доцільно буде використовувати у автоматизованих магазинах та закладах для аналізу залежностей поведінки відвідувачів та їхнього віку. Такі заклади популяризуються за кордоном, та невдовзі з'являться і в Україні.

Розроблену систему буде використано для персоналізації інтерактивної взаємодії людини з високотехнологічною розважальною системою "Робот-бармен". У реальному часі система зможе визначити вік людини за допомогою зображення з камери, що дозволить персоналізувати голосову взаємодію з розважальною системою.

### Аналіз досліджень та публікацій

Запропонована в [1] робота заснована на двійковому класифікаторі, який визначає те, чи два вхідні зображення об'єднані в один і той же клас, і тренує модель згорткової нейронної мережі (CNN), використовуючи метод глибокого метричного навчання на основі сіамської мережі. Для зближення результатів навчальної сіамської мережі два класи, для яких вікові відмінності нижче певного рівня дистанції, розглядаються як один клас, тому відношення позитивних зображень бази даних збільшується. Метод глибокого метричного навчання навчає модель CNN вимірювати схожість лише на основі даних про вік, але ми виявили, що накопичені дані про стать також можна використовувати для порівняння віку. Таким чином, ми застосували багатозадачний підхід до навчання, щоб розглянути гендерні дані для більш точної оцінки віку.

Публікація [2] зосереджується на оцінці віку, класифікації віку та класифікації статі за нерухомими зображеннями обличчя людини. Автори навчають різні моделі для кожної проблеми, а також проводять порівняння між створенням спеціальної архітектури CNN (згорткової нейронної мережі) та використанням різних архітектур CNN як екстракторів функцій, а саме VGG16, попередньо навченого на VGGFace, ResNet50 та SE-ResNet50, попередньо навченого на VGGFace2. Автори також забезпечили базу продуктивності різних алгоритмів машинного навчання для вилучення функцій, що дало нам найкращі результати. Було помічено, що навіть проста лінійна регресія, натренована на таких витягнутих ознаках, перевершувала навчання CNN, ResNet50 і ResNeXt50 з нуля для оцінки віку.

У роботі [3] запропоновано попереднє дослідження щодо підвищення продуктивності попередньо навчених глибоких мереж шляхом застосування стратегій постобробки. Основна перевага щодо стратегій точного налаштування полягає в простоті та низьких обчислювальних витратах на етапі постобробки. Наскільки відомо, ця стаття є першим дослідженням щодо оцінки віку, яке пропонує використовувати стратегії постобробки для функцій, отриманих за допомогою попередньо навчених глибоких мереж. У представленому методі використовується набір попередньо навчених згорткових нейронних мереж (CNN) для вилучення ознак із вхідного зображення обличчя. Потім метод виконує об'єднання на рівні ознак, зменшує розмірність простору ознак і оцінює вік людини за допомогою нейронної мережі з прямим зв'язком (FFNN). Було оцінено ефективність даного методу на загальнодоступному наборі даних (Adience Benchmark of Unfiltered Faces for Gender and Age Classification) і на наборі даних неідеальних зразків, на які впливали контрольовані повороти, які ми зібрали в нашій лабораторії. Цей метод оцінки віку отримав кращі або порівнянні результати щодо найсучасніших методів і досяг задовільної роботи в неідеальних умовах. Результати також показали, що CNN, навчені на загальних наборах даних, можуть отримати задовільну точність для різних типів зображень перевірки, також без застосування методів тонкого налаштування.

Автори [4] намагались вирішити проблему різних оптичних та рухових ефектів розмиття. Такі ефекти зазвичай викликають проблеми з повним охопленням таких рис обличчя, як зморшки, які є важливими для оцінки віку, тим самим погіршуючи точність. Більшість попередніх досліджень щодо оцінки віку було проведено для вхідних зображень, майже без ефекту розмиття. Щоб подолати це обмеження, автори пропонують використовувати глибоку згорткову нейронну мережу ResNet-152 для оцінки віку, яка стійка до різних оптичних ефектів і ефектів розмиття руху від датчиків видимого світла. Було проведено експерименти з різними оптичними та розмитими зображеннями, створеними з баз даних паркової лабораторії старіння розуму (PAL) та баз даних черепно-лицевої поздовжньої морфологічної бази даних обличчя (MORPH), які є загальнодоступними. Згідно з результатами, запропонований метод показав кращу оцінку віку, ніж попередні методи.

Стаття [5] представляє багатоцільовий алгоритм для одночасного виявлення обличчя, вирівнювання обличчя, оцінки пози, розпізнавання статі, виявлення посмішки, оцінки віку та розпізнавання обличчя за допомогою єдиної глибокої згорткової нейронної мережі (CNN). Запропонований метод використовує багатозадачну систему навчання, яка регулює спільні параметри CNN і створює синергію між різними доменами та завданнями. Великі експерименти показують, що мережа краще розуміє обличчя і досягає найсучасніших результатів для більшості цих завдань.

У дослідженні [6] показано, що для оцінки віку та гендерної класифікації можна отримати значне покращення, вивчаючи уявлення з використанням глибоких згорткових нейронних мереж (CNN). Метод нейронної мережі з прямим зв'язком, використаний у цьому дослідженні, підвищує надійність для дуже варіабельних завдань розпізнавання без обмежень для визначення статі та вікової групи. Це дослідження було проаналізовано та підтверджено для прогнозування статі та оцінки віку як для набору даних обличчя Ессекса, так і для контрольного показника Adience. Отримані результати показують, що запропонований підхід забезпечує значний приріст продуктивності, дана модель досягає дуже цікавої ефективності та найсучаснішого рівня оцінки як за віком, так і за статтю.

На основі легкої згорткової нейронної мережі в статті [7] пропонується покращена мережа ShuffleNetV2, заснована на механізмі змішаної уваги (MA-SFV2: Mixed Attention-ShuffleNetV2), шляхом перетворення вихідного шару, об'єднання методів класифікації та регресійної оцінки віку та виділення важливих функцій попередньої обробки зображень і методів збільшення даних. Вплив векторів шуму, таких як інформація про навколишнє середовище, не пов'язана з обличчями на зображенні, зменшується, тому точність остаточної оцінки віку можна порівняти з сучасною.

У [8] автори пропонують нову архітектуру CNN під назвою Fusion Network (Fusion-Net) для вирішення проблеми оцінки віку. Окрім зображення всього обличчя, FusionNet послідовно приймає кілька вікових патчів на обличчі як частину вхідних даних, щоб підкреслити вікові риси. За допомогою експериментів було показано, що FusionNet значно перевершує інші найсучасніші моделі за тестом MORPH II.

У [9] розглядався ефект використання оцінки віку на основі нейронної мережі глибокої згортки (CNN) як зручності для моніторингу пішоходів, які переходять на перехрестях. У цій роботі були використані дві популярні попередньо навчені моделі глибоких згорткових нейронних мереж (DCNN), які нещодавно досягли найкращої продуктивності в задачах виділення рис обличчя: VGG-Face і ResNet-50. Автори об'єднали ці дві моделі для підвищення ефективності запропонованої системи. Було проведено експерименти, щоб оцінити систему на основі набору даних VGGFace2, що складається з 3,31 мільйона зображень обличчя. За результатами експерименту було виявлено розрив у результативності виявлення між цими віковими групами: діти від (00-10) років та люди похилого віку від 55 років і більше. Крім того, було відзначено, що запропонована модель оцінки віку пішохода є високою, а також хороший результат можна отримати, якщо використовувати модель за новим призначенням.

Автори [10] запропонували високоефективну систему оцінки віку із спільною оптимізацією алгоритму оцінки віку та системою глибокого навчання. У співпраці з міською мережею наглядю ця система може забезпечити аналіз вікових груп для розумної демографії. Спочатку було створено тривірневу архітектуру туманних обчислень, що включає край, туман і шар хмари, яка безпосередньо обробляє оцінку віку з необроблених відео. Далі проводиться оптимізація алгоритму оцінки віку на основі CNN з розподілом міток і дивергенцією K-L, вбудованим у шар туману, і оцінюємо модель на основі останнього набору даних про дике старіння. Результати експерименту свідчать, що дана система динамічно збирає демографічні дані на віддаленій відстані без контакту та автоматично проводить аналіз населення міста та навчання вікової моделі було прискорено без втрати прогресу навчання чи якості моделі. Наскільки відомо, це перша інтелектуальна демографічна система, яка має потенційні можливості для підвищення ефективності розумних міст і міського життя.

### Формулювання цілей статті

Метою роботи є створення швидкодіючої та якісної системи апроксимації віку людини з можливістю аналізу зображення та роботи у режимі реального часу для застосування у високотехнологічній розважальній системі.

### Виклад основного матеріалу

Проаналізувавши роботи у численних джерелах, було прийнято рішення для навчання моделей нейронної мережі обрати кілька наборів даних та об'єднати їх частини. У таблиці 1 наведено основні характеристики проаналізованих датасетів.

Таблиця 1 [1]

Огляд наборів даних, що містять обличчя людей з казаним віком

База даних	Кількість прикладів	Вік	Обличчя чоловіків, %	Особливості
UTKFace Dataset	20000	0-116	52	Обличчя різних етносів
IMDB-WIKI	523051	0-100	59	Обличчя відомих людей
All-Age-Faces	13322	2-80	44,6	Переважає азіатська зовнішність

### Етапи процесу апроксимації віку

Автоматизований процес апроксимації віку поділяють на три основні етапи: виявлення обличчя; аналіз зображення з вилученням ознак; класифікація вікової категорії.

**Виявлення обличчя.** Виявлення обличчя на зображенні є складним завданням через велику кількість варіацій ключових характеристик, а саме різні розміри обличчя на зображенні, різні перспективи, освітлення, фон та саму різноманітність людських лиць. Також на зображенні можуть бути й інші об'єкти, окрім людей.

Алгоритми розпізнавання обличчя зазвичай спершу шукають пару людських очей, адже це одна з характеристик, яку найлегше виявити. Знайшовши область обличчя, він виконує додаткові тести, щоб підтвердити, що він насправді виявив обличчя. Одним з найпоширеніших для виявлення обличчя є метод каскад Хаара, що використовує набори прямокутних областей (фільтрів) з різними числовими характеристиками, які переміщуються по зображенню і порівнюються з його частинами для виявлення збігів. Даний метод має високу швидкодію але його недоліком є складність визначення числових характеристик прямокутних областей для ефективного аналізу зображення. Щоб побороти цей недолік використовують згорткові нейронні мережі, що надає змогу автоматично визначити оптимальні значення фільтрів та збільшує ефективність проведеного аналізу, проте вимагає більшу обчислювальну потужність. Після виявлення обличчя на зображенні здійснюється кадрівання, що залишає лише область зображення самого обличчя та певну маржу.

**Аналіз зображення обличчя.** Для проведення аналізу зображення обличчя найоптимальніше використовувати згорткову нейронну мережу, що на відміну від традиційних класифікаторів з наперед визначеними значеннями фільтрів, дозволяє автоматизувати процес визначення ваг фільтрів, що виявляють наявність певних характеристик на зображенні. Це здійснюється за допомогою тренування нейронної мережі на навчальних даних. Проаналізувавши численну кількість робіт, було прийнято рішення обрати кілька архітектур згорткових нейронних мереж та натренувати їх, після чого на основі швидкодії та коректності вибрати найоптимальнішу.

**Класифікація вікової категорії.** Класифікація вікової категорії зображеної людини відбувається на основі аналізу нейронною мережею виявлених за допомогою згорткової нейронної мережі характеристик обличчя та їх відповідності певним віковим категоріям.

Принцип роботи системи наведено на рис. 1.

### Висновки з даного дослідження і перспективи подальших розвідок у даному напрямі

В результаті проведеної роботи було створено систему апроксимації віку людини за зображенням її обличчя. Ця розробка дозволяє визначити вік особи на завантаженому зображенні, а також здійснювати це у реальному часі. Детальне дослідження задачі апроксимації віку людини показала, що система має застосування у багатьох галузях, наприклад розваг, збору даних, безпеки та моніторингу. Оскільки ця галузь недостатньо досліджена в Україні, то дана розробка є доцільною та актуальною. Точність розробленої системи становить 80%, що є доволі хорошим показником, але наступні дослідження можуть бути

спрямовані на досягнення вищої якості за допомогою кращої підготовки даних, збільшення кількості епох при тренуванні нейронної мережі та скрупульознішого підбору гіперпараметрів.

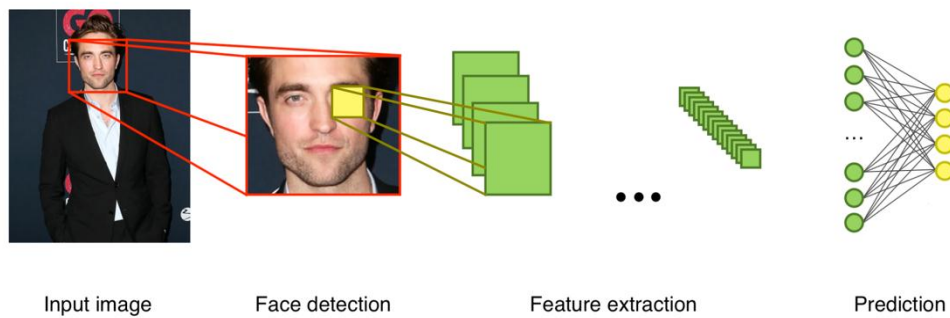


Рис. 1. Принцип роботи системи

### Література

1. Y. Jeong, S. Lee, D. Park, i K. H. Park, «Accurate age estimation using multi-task Siamese network-based deep metric learning for frontal face images», *Symmetry*, 10, 9, 2018, doi: 10.3390/sym10090385.
2. V. Sheoran, S. Joshi, i T. R. Bhayani, «Age and Gender Prediction Using Deep CNNs and Transfer Learning», *Communications in Computer and Information Science*, 1377 CCIS, p. 293–304, 2021, doi: 10.1007/978-981-16-1092-9\_25.
3. Anand, R. D. Labati, A. Genovese, E. Munoz, V. Piuri, i F. Scotti, «Age estimation based on face images and pre-trained convolutional neural networks», 2018, вип. 2018-January, p. 1–7. doi: 10.1109/SSCI.2017.8285381.
4. J. S. Kang, C. S. Kim, Y. W. Lee, S. W. Cho, i K. R. Park, «Age estimation robust to optical and motion blurring by deep residual CNN», *Symmetry*, 10, 4, 2018, doi: 10.3390/sym10040108.
5. R. Ranjan, S. Sankaranarayanan, C. D. Castillo, i R. Chellappa, «An All-In-One Convolutional Neural Network for Face Analysis», 2017, p. 17–24. doi: 10.1109/FG.2017.137.
6. M. K. Benkaddour, «CNN based features extraction for age estimation and gender classification», *Informatica (Slovenia)*, 45, 5, p. 697–703, 2021, doi: 10.31449/INF.V45I5.3262.
7. X. Liu, Y. Zou, H. Kuang, i X. Ma, «Face image age estimation based on data augmentation and lightweight convolutional neural network», *Symmetry*, 12, 1, 2020, doi: 10.3390/SYM12010146.
8. H. Wang, X. Wei, V. Sanchez, i C.-T. Li, «Fusion network for face-based age estimation», 2018, p. 2675–2679. doi: 10.1109/ICIP.2018.8451606.
9. N. Y. Abdullah, M. T. Ghazal, i N. Waisi, «Pedestrian age estimation based on deep learning», *Indonesian Journal of Electrical Engineering and Computer Science*, 22, 3, p. 1548–1555, 2021, doi: 10.11591/ijeecs.v22.i3.pp1548-1555.
10. Z. Hu, P. Sun, i Y. Wen, «Speeding-up age estimation in intelligent demographics system via network optimization», 2018-May. doi: 10.1109/ICC.2018.8422409.