BOOK OF ABSTRACTS

ICAS CONFERENCE

INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

11th International Conference on Advances in Statistics
25 - 27 April 2025

ICAS CONFERENCE

INTERNATIONAL CONFERENCE ON ADVANCES IN STATISTICS

25 - 27 April 2025

Bologna – ITALY

http://www.icasconference.com/

ICAS'2025

11th International Conference on Advances in Statistics

Published by the ICAS Secretariat

Editors:
Prof. Dr. Ismihan BAYRAMOGLU

ICAS Secretariat
Büyükdere Cad. Ecza sok. Pol Center 4/1 Levent-İstanbul
E-mail: icasconference.academic@gmail.com
http://www.icasconference.com

ISBN: 978-605-65509-7-3

INTERNATIONAL SCIENTIFIC COMMITTEE

Prof. Dr. Alexei Stepanov
Immanuel Kant Baltic Federal University – Russia

Prof. Dr. Ayman BAKLEEZI
Qatar University – Qatar

Prof. Dr. Barry C. ARNOLD
University of California, Riverside – USA

Prof. Dr. George YANEV
The University of Texas Rio Grande Valley – USA

Prof. Dr. Hamparsum BOZDOGAN
The University of Tennessee – USA

Prof. Dr. Hamzeh TORABI
Yazd University – IRAN

Prof. Dr. Ilham AKHUNDOV
University of Waterloo – Canada

Prof. Dr. Jorge NAVARRO
Universidad de Murcia – Spain

Prof. Dr. Jose Maria SARABIA
University of Cantabria – Spain

Prof. Dr. Leda MINKOVA
University of Sofia – Bulgaria

Prof. Dr. Narayanaswamy BALAKRISHNAN
McMaster University – Canada

Prof. Dr. Nikolai KOLEV
University of Sao Paulo – Brasil

Prof. Dr. Sarjinder SINGH
Texas A&M University-Kingsville – USA

Prof. Dr. Stelios PSARAKIS
Athens University of Economics & Finance – Greece

Prof. Valerii B. Nevzorov
St. Petersburg State University – Russia

LOCAL SCIENTIFIC COMMITTEE

Prof. Dr. Aydin ERAR
Mimar Sinan Fine Arts University – Turkey

Prof. Dr. Baris ASIKGIL
Mimar Sinan Fine Arts University – Turkey

Prof. Dr. Coskun KUS
Selcuk University – Turkey

Prof. Dr. I. Esen YILDIRIM
Marmara University – Turkey

Prof. Dr. Esra AKDENIZ
Marmara University – Turkey

Prof. Dr. Gulay BASARIR
Mimar Sinan Fine Arts University – Turkey

Prof. Dr. Ismail KINACI
Selcuk University – Turkey

Prof. Dr. Mehmet Fedai KAYA
Selcuk University – Turkey

Prof. Dr. Mujgan TEZ
Marmara University – Turkey

Prof. Dr. Sahamet BULBUL
Fenerbahce University – Turkey

Prof. Dr. Yazgi TUTUNCU
Izmir University of Economics, Turkey

Assoc. Prof. Dr. Konul KAVLAK
Wageningen University – Holland

ORGANIZATION COMMITTEE

Prof. Dr. Ismihan BAYRAMOGLU
Izmir University of Economics – Turkey
Conference Chair

Prof. Dr. Fatma NOYAN TEKELI
Yıldız Technical University – Turkey

Prof. Dr. Gulhayat GOLBASI SIMSEK
Yıldız Technical University – Turkey

Prof. Dr. Seher Nur SULKU
Ankara Hacı Bayram Veli University – Turkey

Assoc. Prof. Dr. Aysegul EREM
Cyprus International University – TRNC

Phd. Ayse BELER
Izmir University of Economics, Turkey

Gokce OZALTUN
Izmir University of Economics, Turkey

Dear Colleagues,

On behalf of the Organizing Committee, I am pleased to invite you to participate in **11th International Conference on Advances in Statistics** which will held in Bologna–ITALY on the dates between 25 – 27 April 2025

All informations are available in conference web site. For more information please do not hesitate to contact us. info@icascconference.com

- Applied Statistics
- Bayesian Statistics
- Big Data Analytics
- Biostatistics
- Comutational Statistics
- Data Analysis and Modeling
- Data Envelopment Analysis
- Data Management and Decision Support Systems
- Data Mining
- Econometrics
- Energy and Statistics
- Enviromental Statictics
- Entrepreneurship
- Mathematical Statistics
- Multivariate Statistics
- Neural Networks and Statistics
- Non-parametric Statistics
- Numerical Methods
- Operations Research
- Optimization Methods in Statistics
- Order Statistics
- Panel Data Modelling and Analysis
- Performance Analysis in Administrative Process
- Philosophy of Statistics
- Public Opinion and Market Research
- Quality Control
- Reliability Theory
- Sampling Theory
- Simulation Techniques
- Spatial Analysis
- Statistical Software
- Statistical Training
- Statistics Education
- Statistics in Social Sciences
- Stochastic Processes
- Supply Chain
- Survey Research Methodology
- Survival Analysis
- Time Series
- Water and Statistics
- Other Statistical Methods

Selected papers will be published in **Journal of the Turkish Statistical Association**. https://dergipark.org.tr/en/pub/ijtsa

We hope that the conference will provide opportunities for participants to exchange and discuss new ideas and establish research relations for future scientific collaborations.

Conference Website : https://icasconference.com
E Mail: icasconference.academic@gmail.com

On behalf of Organizing Committee:
Conference Chair
Prof. Dr. İsmihan BAYRAMOĞLU
Izmir University of Economics

25 APRIL 2025 FRIDAY

BERLIN MEETING ROOM

Welcome Speech / 09:00 – 09:10
Prof. Dr. Ismihan BAYRAMOGLU / Conference Chair
Izmir University of Economics , Izmir, Turkiye

#1 Keynote Speech / 09:10 – 09:50
Prof. Dr. Barry C. Arnold
University of California, Riverside, USA

Speech Title: Conditionally Specified Bivariate Models: Old and New

SET UP BREAK 09:50 – 10:00

SESSION A (10:00 – 11:20)

SESSION CHAIR: Prof. Dr. Aysen D. AKKAYA

BERLIN MEETING ROOM

10:00 – 10:20
PAPER TITLE : Random Effect Functional Cox Regression Model
AUTHOR(S) : **Deniz INAN**, Ufuk BEYAZTAS, Carmen D. TEKWE, Xiwei CHEN, Roger S. ZOH

10:20 – 10:40
PAPER TITLE : A Deep Learning Approach to Spatial Functional Regression with Mixed Covariates
AUTHOR(S) : Merve BAŞARAN, **Ufuk BEYAZTAŞ**

10:40 – 11:00
PAPER TITLE : Forecasting the Process of Screened / Randomized Patients and Optimal Early Stopping Time in Multicentre Clinical Trials
AUTHOR(S) : **Volodymyr ANISIMOV**

11:00 – 11:20
PAPER TITLE : Robust Inference for a Fat-Tailed Skew Distribution
AUTHOR(S) : **Ayşen D. AKKAYA**, M. Qamarul ISLAM

TEA & COFFEE BREAK 11:20 – 11:40

SESSION B (11:40 – 13:00)

SESSION CHAIR: Prof. Dr. Deniz INAN

BERLIN MEETING ROOM

11:40 – 12:00

PAPER TITLE : Enhanced Robust and Efficient Estimation Methods for Panel Data Regression with Random Effects

AUTHOR(S) : **Beste Hamiye BEYAZTAŞ**, Abhijit MANDAL, Soutir BANDYOPADHYAY

12:00 – 12:20

PAPER TITLE : Spatial Autoregressive Quantile Regression for Scalar-on-Function Mode

AUTHOR(S) : **Müge MUTİŞ**, Ufuk BEYAZTAŞ, Filiz KARAMAN, Han Lin SHANG

12:20 – 12:40

PAPER TITLE : Reliability of Complex Systems Consisting of Two Types of Mixed Components

AUTHOR(S) : **Hümay Zehra ÖZER**, İsmihan BAYRAMOĞLU

12:40 – 13:00

PAPER TITLE : The Reliability of the Complex System Having Active and Redundant Components

AUTHOR(S) : **Könül BAYRAMOĞLU KAVLAK**

LUNCH BREAK 13:00 – 13:30

#2 Keynote Speech / 13:30 – 14:00

Prof. Dr. Umberto Cherubini

University of Bologna, ITALY

Betting Around the Clock: Long-Term Investment with Time Change

14:10 Departure from Conference Venue for Florence Tour and Conference Dinner

Please check **conference social program** for more information

23:30 Back to Hotel After Tour & Conference Dinner

Please check **conference social program** for more information

CONFERENCE VENUE

ZAN HOTEL EUROPA BOLOGNA, ITALY

26 APRIL 2025 SATURDAY

SESSION C (09:00 – 10:20)
SESSION CHAIR: Prof. Dr. Michail D. TODOROV
BERLIN MEETING ROOM

09:00 – 09:20
PAPER TITLE : Evaluation of Provinces According to Losses in Businesses After February 6 Earthquakes Using Multi-Criteria Decision-Making Methods
AUTHOR(S) : **Pelin TOKTAŞ**

09:20 – 09:40
PAPER TITLE : A Multicriteria Optimization Approach for Investment Project Financing
AUTHOR(S) : Borislav CHAKAROV, **Meglena LAZAROVA**, Ivan GEORGIEV, Slavi GEORGIEV

09:40 – 10:00
PAPER TITLE : Multi-Level Thresholding Image Segmentation Using Metaheuristic Algorithm
AUTHOR(S) : Eda ÖZKUL, **Muharrem KORKMAZ**

10:00 – 10:20
PAPER TITLE : The Role of Banking Heterogeneity in the Relationship Between Financial Development and Entrepreneurship
AUTHOR(S) : Cristian BARRA, **Christian D'ANIELLO**

TEA & COFFEE BREAK 10:20 – 10:40

#3 Keynote Speech / 10:40 – 11:10

Prof. Dr. Jorge NAVARRO
University of Murcia, SPAIN
Stochastic Properties of Sums of Dependent Random Variables (Risks)

SET UP BREAK 11:10 – 11:20

SESSION D (11:20 – 12:40)

SESSION CHAIR: Dr. Volodymyr ANISIMOV

BERLIN MEETING ROOM

11:20 – 11:40
PAPER TITLE : A Note on a Class of Estimates for the Regression Function
AUTHOR(S) : Mounir ARFI

11:40 – 12:00
PAPER TITLE : Use of Linear Dynamic Models for Time Series Prediction
AUTHOR(S) : Lubos MAREK, Richard HINDLS, **Stanislava HRONOVA**

12:00 – 12:20
PAPER TITLE : A Skew Marshall-Olkin Distribution
AUTHOR(S) : Leda Minkova

12:20 – 12:40
PAPER TITLE : A Fuzzy C-Means Image Segmentation Algorithm Using Metaheuristics
AUTHOR(S) : Özge TEZEL, **Onur OKUYUCU**

LUNCH BREAK 12:40 – 13:20

#4 Keynote Speech / 13:20 – 13:50

Anna Dembińska
Warsaw University of Technology, POLAND
Predicting Missing Data Using Inactivity Times of Components Upon System Failure

SET UP BREAK 13:50 – 14:00

SESSION E (14:00 – 15:40)

SESSION CHAIR: Prof. Dr. Leda MINKOVA

BERLIN MEETING ROOM

14:00 – 14:20
PAPER TITLE : Using HMD Methods Protocol to Estimate Turkish Population at the Advanced Ages
AUTHOR(S) : **Özer BAKAR**, Murat BÜYÜKYAZICI, İsmet KOÇ

14:20 – 14:40
PAPER TITLE : On a Splitting Method for Solving the Nonlinear Schrödinger Equation and Its Generalization to the Manakov System
AUTHOR(S) : **Michail TODOROV**, Meglena LAZAROVA,

14:40 – 15:00
PAPER TITLE : Development of a Deep Learning-Based Hybrid Model for Diagnosing Lung Cancer Using CT Scan Images
AUTHOR(S) : **Yasemin GÜNTER** , Ayşenur ÖZTÜRK, Fatma NOYAN TEKELİ

15:00 – 15:20
PAPER TITLE : Airport Customer Segment Ranking: Hybrid TOPSIS Approach
AUTHOR(S) : **Beyza AHLATÇIOĞLU ÖZKÖK**, Serhan KARADENİZ, Sezin ÖZTÜRK

15:20 – 15:40
PAPER TITLE : The Integration of Artificial Intelligence Tools in Teaching Statistics
AUTHOR(S) : **Laura ASANDULUI**, Christiana SANDU

TEA & COFFEE BREAK 15:40 – 16:00

SESSION F (16:00 – 17:30)

SESSION CHAIR: Prof. Dr. Ismihan BAYRAMOGLU

BERLIN MEETING ROOM

16:00 – 16:20
PAPER TITLE : Smart Classrooms of the Future: AI-Driven Support for Teachers and Students
AUTHOR(S) : **Laura PAGANI**, Paolo GEMELLI

16:20 – 16:40
PAPER TITLE : Hybrid Waste Management System: Sensor and Vision-Based Data Collection and Artificial Intelligence-Based Classification
AUTHOR(S) : Bedirhan Gürsu,Berkant Özkan, Eren Okumus, Ömer Faruk Gülhan, **Özden Niyaz**

16:40 – 17:00
PAPER TITLE : Simulation Framework for Real-Time Management and Control of Manufacturing Systems
AUTHOR(S) : **Iwona LAPUNKA**, Piotr WITTBRODT

17:00 – 17:20
PAPER TITLE : Lung and Colon Cancer Histopathological Image Classification Using Deep Learning Approaches
AUTHOR(S) : Oussama Bouguerra**, Youcef Brik,** Bilal Attallah

17:20 – 17:30 (CLOSING SPEECH)
: Prof. Dr. Ismihan BAYRAMOGLU, Izmir University of Economics
Conference Chair

END OF CONFERENCE

27 APRIL 2025 SUNDAY

PISA & LUCCA CITY TOUR WITH GUIDE

Please check **conference social program** for more informatio

Contents

Conditionally specified bivariate models: old and new

Barry C. Arnold and Ashis SenGupta
University of California, Riverside, USA

Abstract:

A sample of old and new examples of conditionally specified bivariate models will be discussed. Cases with conditionals in exponential families are well-known. Some related parallel constructions are possible in non-exponential family settings. In all cases the Stephanos-Suto-Levi-Civita functional equation plays a key role.

RANDOM EFFECT FUNCTIONAL COX REGRESSION MODEL

Deniz INAN[1], Ufuk BEYAZTAS[2], Carmen D. TEKWE[3], Xiwei CHEN[3], Roger S. ZOH[3]

[1]Department of Statistics, Marmara University, Turkey e-mail: denizlukuslu@marmara.edu.tr

[2]Department of Statistics, Marmara University, Turkey e-mail: ufuk.beyaztas@marmara.edu.tr

[3]Department of Epidemiology and Biostatistics, Indiana University, School of Public Health email: ctekwe@iu.edu

[3]Department of Epidemiology and Biostatistics, Indiana University, School of Public Health email: xiweichen@iu.edu

[3]Department of Epidemiology and Biostatistics, Indiana University, School of Public Health email: rszoh@iu.edu

Abstract

we introduce a functional linear Cox regression model with frailty to address unobserved heterogeneity in survival data involving functional covariates. Although conventional Cox models and their functional extensions are widely used, they are limited in handling frailty effects or random components for capturing individual differences not explained by observed covariates. Our proposed model integrates both scalar and functional covariates with a frailty term to account for unmeasured influences, thereby providing a robust framework for survival analysis in high-dimensional settings. We derive parameter estimates using functional principal component analysis and employ penalized partial likelihood to accommodate the frailty structure. A comprehensive simulation study demonstrates the model's superior performance in terms of estimation accuracy and predictive capability over traditional functional linear Cox regression models and penalized functional regression models, especially under high-frailty conditions. In addition, we apply the proposed method to the analysis of a dataset from the National Health and Nutrition Examination Survey, revealing significant associations between physical activity patterns and mortality, even in frail subpopulations. Our results underscore the effectiveness of the proposed method in accurately capturing the complexity of survival data with both functional predictors and frailty, highlighting its potential applications in biomedical research where unobserved heterogeneity plays a critical role. Our method has been implemented in the *fcoxFr* R package.

Key Words: *Cox regression; Functional principal component analysis; Frailty; Model selection; Smoothing*

A DEEP LEARNING APPROACH TO SPATIAL FUNCTIONAL REGRESSION WITH MIXED COVARIATES

Merve BASARAN[1], Ufuk BEYAZTAS[1],

[1]Marmara University, Istanbul, Türkiye basaran.merve@icloud.com

[1]Marmara University, Istanbul, Türkiye ufuk.beyaztas@marmara.edu.tr

Abstract

This paper introduces a novel spatial functional deep learning framework for modeling scalar outcomes using both scalar and functional predictors while accounting for spatial dependence. The proposed approach consists of two key stages: first, the spatial autocorrelation parameter is estimated using a functional principal component-based spatial scalar-on-function regression model; second, the scalar outcome is predicted using a functional deep neural network, where the first layer, termed the functional layer, efficiently handles functional covariates. To assess the effectiveness of the proposed method, we conduct a series of Monte Carlo simulations under both linear and nonlinear settings, comparing our approach against a non-spatial functional deep neural network and a linear spatial functional regression model. Additionally, we evaluate the predictive performance of the proposed method on an empirical dataset. Our findings indicate that the proposed spatial functional deep learning approach outperforms existing methods when spatial dependence is moderate to strong, particularly in scenarios involving nonlinear relationships. Furthermore, the method remains competitive even in cases of weak spatial dependence and linear relationships.

Key Words: *Functional data; Functional principal component analysis; Neural network; Prediction; Spatial dependence*

Acknowledgements

This research was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) (grant no. 124F096).

FORECASTING THE PROCESS OF SCREENED/RANDOMIZED PATIENTS AND OPTIMAL EARLY STOPPING TIME IN MILTICENTRE CLINICAL TRIALS

Volodymyr ANISIMOV[1]

[1]Data Science, Center for Design & Analysis, Amgen, London, UK; vanisimo@amgen.com

Abstract

In randomized multicentre clinical trials, the patients typically undergo a screening process before being randomized into the trial with some probability. The success of a trial often depends on accurately forecasting the time required to archive the target sample size while minimizing over-sampling.

A novel analytic methodology for jointly modeling patient screening and randomization processes is introduced. The arrival of patients for screening is modelled using a Poisson-gamma model developed in the author's earlier work [1.2], where the flows of patient arrival at different clinical centres follow doubly stochastic Poisson processes with the rates treated as independent gamma distributed random variables. Center initiation times are accounted for as fixed or random variables. The subsequent randomization of patients after completion screening, considering their probability of randomization, is described using hierarchic binomial or beta-binomial models.

To forecast at interim time the remaining number of patients to be randomized and time to stop trial, first the process of patients arriving for screening is investigated. The parameters of a Poisson-gamma model describing this process are estimated using maximum likelihood technique proposed in [1,2]. The posterior rates in individual centres for this model also have gamma distribution with parameters calculated using Bayesian adjustment and individual centre data (the number of recruited patients and the recruitment window).

For predicting the process of randomized patients out of the patients in screening, a new technique is proposed leveraging an approximative Poisson-gamma process. This approximation uses the approach developed in [3]. Additionally, an innovative optimal stopping rule is derived. It is proposed to stop recruitment prior to reaching a sample size accounting for predictive number of randomized patients among those still in the screening pipeline to minimize over-sampling. Analytic expressions for the optimal stopping rule are provided accounting also for potential errors in estimating the probability of randomization. The developed methodology is validated through R-based simulations.

Key Words: *Multicentre clinical trial; Screening/randomization process; Poisson-gamma recruitment model; Forecasting; Optimal stopping time*

References

[1] Anisimov V, Fedorov V (2007) Modelling, prediction and adaptive adjustment of recruitment in multicentre trials, Statistics in Medicine, 26, 4958-4975.

[2] Anisimov V (2020) Modern analytic techniques for predictive modeling of clinical trial operations, chapter 8, in book “Quantitative Methods in Pharmaceutical Research and Development: Concepts and Applications”, Marchenko O.V., Katenka N.V., (edt), Springer, Cham**,** 361-408.

[3] Anisimov V, Austin M (2020) Centralized statistical monitoring of clinical trial enrollment performance, Communications in Statistics - Case Studies and Data Analysis, v. 6, iss. 4, 392-410.

ROBUST INFERENCE FOR A FAT-TAILED SKEW DISTRIBUTION

Aysen D. AKKAYA[1], M. Qamarul ISLAM[2],

[1] Department of Statistics, METU, 06800 Ankara, Turkey: akkay@metu.edu.tr

[2] Department of Statistics, METU, 06800 Ankara, Turkey: mislam@metu.edu.tr

Abstract

Non-normal behavior of data distorts the linearity of the likelihood methodology and, hence, the resulting maximum likelihood (ML) estimators cannot be obtained in explicit form. The use of modified maximum likelihood (MML) method resolves this problem and provides robust and explicit estimators for the parameters of interest (e.g. location and scale) [1]. However, its use in machine learning-based studies is limited due to the requirement that the values of the involved shape parameters should be known. In the case of unknown shape parameters, an adaptive modified maximum likelihood (AMML) method has been proposed in the literature [2, 3], which combines the MML method with the well-known Huber inference methods. The non-normality considered here can be expressed in terms of asymmetry and fat-tailedness of the relevant distribution. In this study, the estimators are developed by using the AMML method for a skew-t distribution. This distribution has tails fatter than a normal distribution and it also exhibits skewness, thus making it suitable to model data depicting non-normal behavior which is commonly observed in various fields of research. Efficiency and robustness of AMML estimators are investigated by extensive simulation study and their properties are compared with least squares (LS) and MML estimators. Relevant tests of hypothesis are developed and explored for their desirable properties. We also provide several applications where the use of such distribution is justified in terms of meaningful statistical hypotheses.

Key Words: *Adaptive modified maximum likelihood estimator; Least square estimator; Modified maximum likelihood estimator; Skew-t distribution; Test of hypothesis.*

References

[1] Tiku ML, Akkaya AD (2004) Robust Estimation and Hypothesis Testing. New Age International Publishers (P): New Delhi.

[2] Donmez A (2010) Adaptive estimation and hypothesis testing methods, PhD thesis, METU, Ankara, Turkey.

[3] Yentur, B, Akkaya, AD, Bayrak, ÖT(2022) Adaptive estimation of autoregression models under long-tailed symmetric distribution. Communications in Statistics - Simulation and Computation 53 (7): 3395-3417.

ENHANCED ROBUST AND EFFICIENT ESTIMATION METHODS FOR PANEL DATA REGRESSION WITH RANDOM EFFECTS

Beste Hamiye BEYAZTAS[1], Abhijit MANDAL[2], Soutir BANDYOPADHYAY[3]

[1]İstanbul Medeniyet University e-mail: beste.sertdemir@medeniyet.edu.tr

[2]The University of Texas et El Paso e-mail: amandal@utep.edu

[3]Colorado School of Mines e-mail: sbandyopadhyay@mines.edu

Abstract

Panel data regression models play a crucial role in various disciplines, including economics, social sciences, and medical research, due to their ability to capture individual-specific heterogeneity and inter-temporal dynamics. Despite their advantages, traditional estimation techniques, particularly least-squares-based methods, are highly sensitive to outliers and data contamination, which can lead to biased parameter estimates and unreliable statistical inference. Addressing this limitation, we propose a robust estimation approach for linear panel data models with random effects based on the **density power divergence (DPD)** criterion. Our approach leverages the **minimum density power divergence estimator (MDPDE)** to achieve a balance between efficiency and robustness by incorporating a tuning parameter that controls the trade-off between robustness and efficiency. In contrast to traditional approaches, which rely on strict model assumptions such as normality and homoscedasticity of errors, our proposed estimator is developed to reduce the adverse effects of different types of outliers, including block-concentrated outliers that tend to cluster within specific cross-sectional units. The asymptotic properties of the proposed estimator and its robustness are thoroughly illustrated. Furthermore, we evaluate its finite-sample performance through extensive simulation studies and an empirical analysis of real-world climate data. Our findings suggest that the proposed method outperforms traditional and existing robust methods, particularly in the presence of contaminated data, ensuring robust and efficient inference in empirical research.

Key Words: *Panel data; Robust estimation; Least squares; Random effects; Density power divergence*

SPATIAL AUTOREGRESSIVE QUANTILE REGRESSION FOR SCALAR-ON-FUNCTION MODEL

Müge MUTİŞ[1], Ufuk BEYAZTAŞ[2], Filiz KARAMAN[3], Han Lin SHANG[4]

[1]Yildiz Technical University, Istanbul, Türkiye; muge.mutis@yildiz.edu.tr

[2]Marmara University, Istanbul, Türkiye; ufuk.beyaztas@marmara.edu.tr

[3]Yildiz Technical University, Istanbul, Türkiye; fkaraman@yildiz.edu.tr

[4]Macquarie University, Sydney, Australia; hanlin.shang@mq.edu.au

Abstract

Spatial autoregressive scalar-on-function quantile regression is a powerful regression model for characterizing the entire conditional distribution of a scalar response variable. It achieves this by considering the spatial dependence structure as an additional estimator alongside a specific functional regressor for data sets containing spatially dependent units. Most existing approaches based on functional quantile regression estimators may often compromise the consistency of the estimators for the parameters of interest in the estimation of this model. This arises from their ignore of the spatial dependence structure. In this context, we present two innovative estimation procedures designed to accurately and efficiently estimate the spatial autocorrelation parameter and the regression coefficient function by taking into account spatial dependency in the spatial autoregressive scalar-on-function quantile regression model. Our procedures project the functional regressor, which belongs to an infinite-dimensional space, onto a finite dimensional space spanned by orthogonal eigenfunction bases using functional principal component analysis. Subsequently, spatial dependence is improved through a set of instrumental variables in the approximate model obtained by replacing the functional regressor with its projections onto these bases. As a result, the final estimators are obtained based on this refined approximate model. The estimation performance of the proposed procedures is evaluated using a series of Monte Carlo experiments and an empirical dataset. The results indicate that the proposed procedures demonstrate favorable performance compared to existing methods.

Acknowledgments

This research was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) (grant no. 124F096).

Key Words: *Functional principal component analysis; Instrumental variables regression; Quantile regression; Scalar-on-function regression model; Spatial dependence.*

Reliability of Complex Systems Consisting of Two Types of Mixed Components

Hümay Zehra ÖZER[1], Prof. Dr. İsmihan BAYRAMOĞLU[2]

[1]humay.ozer@std.izmirekonomi.edu.tr

[2]ismihan.bayramoglu@izmirekonomi.edu.tr

Abstract

We consider a coherent system consisting of two types of randomly selected components. The first type of components have lifetimes $X_1, X_2, \ldots, X_{n_1}$ with joint distribution function $F(x_1, x_2, \ldots, x_{n_1})$, and the second type of components have lifetimes $Y_1, Y_2, \ldots, Y_{n_2}$ with joint distribution function $G(y_1, y_2, \ldots, y_{n_2})$.

The components' lifetimes of this system are $W_1, W_2, \ldots, W_n$, where $n = n_1 + n_2$ and W_i is either X_k, $k \in \{1,2,\ldots,n_1\}$ or Y_j, $j \in \{1,2,\ldots,n_2\}$.

The reliability of the system is studied for particular multivariate distributions under the condition of exchangeability of random variables $X_1, X_2, \ldots, X_{n_1}$ and $Y_1, Y_2, \ldots, Y_{n_2}$. Some numerical calculations in tables and some illustrative graphs are provided.

Key Words: *Exchangeable Random Variables, Coherent Systems Reliability*

References

[1] M. Asadi, I. Bairamov, On the mean residual life function of the k-out-of-n systems at system level, IEEE-Transactions on Reliability, vol. 55, pp. 314–318, 2006.

[2] M. Asadi, I. Bairamov, A note on the mean residual life function of parallel systems, Communications in Statistics — Theory and Methods, vol. 34, pp. 1–12, 2005.

[3] S. Gurler, I. Bairamov, Parallel and k-out-of-n: G systems with nonidentical components and their mean residual life functions, Applied Mathematical Modelling, vol. 33, no. 2, pp. 1116–1125, 2009.

[4] Bairamov, I. and Parsi, S. (2011) Order statistics from mixed exchangeable random variables. Journal of Computational and Applied Mathematics, Volume 35, Issue 16, 4629-4638.

[5] I. Bayramoglu (Bairamov), S. Eryılmaz, Order statistics of dependent sequences consisting of two differentsets of exchangeable variables, Journal of Computational and Applied Mathematics, vol. 286, pp. 1–6, 2015

The reliability of the complex system having active and redundant components

Konul BAYRAMOGLU KAVLAK[1]

[1]Wageningen University and Research

konul.bayramoglukavlak@wur.nl

Abstract

The rapid advancement of modern technologies has amplified the need for robust and reliable system designs, particularly in machinery where component failure can lead to costly downtime or hazards. In this study, we analyze a complex coherent system composed of one active component and N−1 redundant components, where each component includes two dependent subcomponents. Assuming dependency between subcomponents introduces a realistic layer of complexity often observed in practical applications. The system under consideration is coherent and monotonic, implying that system performance improves as more components remain functional. We derive the system's reliability function and examine the impact of redundancy in the presence of subcomponent dependence. Furthermore, we conduct a cost analysis based on reliability and probabilistic considerations, offering insights into optimal system design from both performance and economic perspectives.

Key Words: *Order statistics, distributions of sums of random vectors, reliability function, cost function*

References

[1] Asadi, M. and Bairamov, I. (2005) A note on the mean residual lifefunction of a parallel system. Communications in Statistics- Theory and Methods, 34, 1–12.

[2] Asadi, M. and Bairamov, I. (2006) On the mean residual life functionof a k-out-of-n structure at the system level. IEEE Transactions on Reliability, 55, 314–318.

[3] Bairamov, I., Ahsanullah, M. and Akhundov, I. (2002) A residual life function of a system having parallel or series structures. Journal of Statistical Theory and Applications 1 (2) (2002), 119–132.

[4] Bairamov, I. and Arnold, B.C. (2008) On the residual lifelengths of the remaining components in an n- k+ 1 out of n system. Statistics & Probability Letters, 78 (8), 945–952.

[5] Bayramoglu, I. (2013) Reliability and mean residual life of complex systems with two dependent components per element. IEEE Transactions on Reliability, 62 (1), 276–285.

[6] Dembinska, A., Nikolov, N. and Stoimenova, RE. (2021) Reliability properties of k-out-of-n systems with one cold standby unit. Journal of Computational and Applied Mathematics, 388,113289.

[7] Gurler, S. and Bairamov I. (2009) Parallel and k-out-of-n: G systems with nonidentical components and their mean residual life functions. Applied Mathematical Mdeling, 33, 116-1125.

[8] Gurler, S. (2013) The Mean Remaining Strength Of Systems In A Stress-Strength Model. Hacettepe Journal of Mathematics and Statistics, 42(2), 181-187.

[9] Eryilmaz, S. and K. Bayramoglu, K. (2012) Residual lifetime of consecu-tive k-out-of-n systems under double monitoring. IEEE Transactions on Reliability, 61 (2012), 792–797.

[10] Eryilmaz , S. (2013) Reliability of a k-out-of-n system equipped with a single warm standby component. IEEE Transactions on Reliability, 55, Issue: 2, 499-503.

[11] Eryilmaz , S. and Tank, F. (2012) On reliability analysis of a two-dependent-unit series system with a standby unit, Appl Math Comput. 218(15):7792–7797

[12] Li, X. and Zhang, Z. (2008) Applied Stochastic Models in Buisiness and Industry, 24, 541–549.

[13] Kavlak, K. (2017) Reliability and Mean Residual Life Functions of Coherent Systems in an Active Redundancy. Naval Research Logistics, 64, 19-28.

[14] Khaledi, B.E, and Shaked, M. (2007) Ordering conditional lifetimes of coherent systems. Journal of Statistical Planning and Inference, 4, 1173-1184.

[15] Zhang, Z. (2010) Ordering conditional general coherent systems with exchangeable components. Journal of Statistical Planning and Inference, Volume 140, Issue 2, 454-460.

[16] Ucer, B. and Gurler, S. (2012) On the mean residual lifetime at system level in two-component parallel systems for the FGM distribution. Hacettepe Journal of Mathematical Statistics, 41, 139–145.

Betting Around the Clock: Long Term Investment with Time Change

Umberto Cherubini

University of Bologna ,umberto.cherubini@unibo.it

Abstract:

We review the longstanding debate between information theorists and financial economics on the asymptotic distribution of optimal investment. We show that maximization of the log of wealth, that is commonly known as Kelly rule, may be actually impossible to reach in a general setting in which the dynamics of the one period returns are represented a process with stochastic time. As it is well known, this covers all the cases in which returns are semi-martingale processes. The topic is illustrated using the Variance Gamma model, with the specific functional shape that is currently used as the representation of the dynamics of asset prices in state of the art option pricing theory.

EVALUATION OF PROVINCES ACCORDING TO LOSSES IN BUSINESSES AFTER FEBRUARY 6 EARTHQUAKES USING MULTI-CRITERIA DECISION-MAKING METHODS

Pelin TOKTAŞ

Başkent University, Faculty of Engineering, Industrial Engineering Department, Ankara, 06790, Türkiye

e-mail: ptoktas@baskent.edu.tr

Abstract

Two catastrophic earthquakes of magnitudes 7.7 and 7.6 struck Turkey on February 6, 2023. These earthquakes, among the most devastating in Turkey's history, struck 11 cities and caused significant property damage, economic disruption, and fatalities. Reports prepared on damage assessments and economic loss estimations carried out as quickly as possible after the earthquake have gained importance.

For the International Labour Organization (ILO), the Adhoc Research Company conducted a study in 2024. To evaluate the current situation of businesses and their recovery efforts in earthquake-affected areas and examine the economic impact on the labour market, the "Enterprises Survey in Post-Disaster Setting" study was started and reported on March 2024. Due to regional differences, cities were affected by the earthquake to different degrees. Except for Gaziantep, where the scope of reconstruction was significantly limited, the analysed data came from the responses of enterprises in Hatay, Kahramanmaraş, Adıyaman, and Malatya. In this report, data from 579 enterprises in the mentioned provinces were included in the evaluation of the impact of the earthquake.

In this study, it is aimed to rank the four provinces affected by the earthquake according to losses in business. While determining the situation of the provinces in terms of job losses after the earthquake, eleven criteria were considered from the report. The Kemeny Median Rank Accordance (KEMIRA-M) method, one of the multi-criteria decision-making methods, was used to rank the provinces. In this method, the criteria are divided into two groups. This study considered eleven criteria taken from the report on job losses in the provinces after the earthquake. Five criteria (loss of life among business employees, permanent physical disability among business employees, physical damage to business real estate, loss/damage to production tools/equipment, and reduction in talent/candidate pool) were examined under the title "Human and Operational Losses." Six criteria (product losses, market losses, supplier losses, customer losses, raw material losses, and distributor losses) were examined under "Commercial and Economic Losses." Then, the ranking that minimizes the difference between the criteria priority rankings determined by three experts was determined as the median priority component (MPC). The RANking COMparison (RANCOM) method was also used for criteria weighting according to the MPC. The consistency of the weights obtained was checked with triad consistency. Then, the weight pair sets for the two criteria groups were used to find the minimum total absolute weighted average differences and the final ranking of the provinces.

As a result of this study, knowing the rankings of the provinces affected by the earthquake according to losses in business will guide the decision on which province to start planning for revitalizing the labour market and economic structures.

Key Words: *KEMIRA-M; RANCOM; Triad Consistency; February 6 Earthquake.*

References

[1] “Enterprises Survey In Post-Disaster Setting“ for ILO, (March 2024)

https://www.ilo.org/sites/default/files/2024-10/Enterprises%20Survey%20in%20Post-Disaster%20Setting%20V02.pdf Last access date: 08/04/2025.

[2] Krylovas, A., Kazimieras Zavadskas E., Kosareva, N. (2016). Multiple Criteria Decision-Making KEMIRA-M Method for Solution of Location Alternatives. *Economic research-Ekonomska istraživanja*, 29(1), 50-65.

[3] Krylovas, A., Kosareva, N., Zavadskas, E.K. (2016). Statistical Analysis of KEMIRA Type Weights Balancing Methods. *Romanian Journal of Economic Forecasting*, 19(3), 19-39.

[4] Krylovas, A., Zavadskas, E. K., Kosareva, N. & Dadelo, S. (2014). New KEMIRA Method for Determining Criteria Priority and Weights in Solving MCDM Problem. *International Journal of Information Technology & Decision Making, 13*, 1119–1133.

[5] Sałabun, W., Shekhovtsov, A., Kizielewicz, B., (2021). A new consistency coefficient in the multi-criteria decision analysis domain. In: Computational Science–ICCS 2021: 21st International Conference, Krakow, Poland, June 16–18, 2021, Proceedings, Part I. Springer, pp. 715–727.

[6] Więckowski, J., Kizielewicz, B., Shekhovtsov, A., & Sałabun, W. (2023). RANCOM: A novel approach to identifying criteria relevance based on inaccuracy expert judgments. *Engineering Applications of Artificial Intelligence*, *122*, 106114.

A Multicriteria Optimization Approach for Investment Project Financing

Borislav Chakarov[1][0009-0009-7220-9771], Meglena Lazarova[2][0009-0001-8723-0014], Ivan Georgiev[3,1][0000-0001-6275-6557] and Slavi Georgiev[3,1][0000-0001-9826-9603]

[1]Faculty of Natural Sciences and Education, University of Ruse, 7004 Ruse, 8 Studentska Str., Bulgaria, bchakarov@uni-ruse.bg, irgeorgiev@uni-ruse.bg, sggeorgiev@uni-ruse.bg

[2]Faculty of Applied Mathematics and Informatics, Technical University of Sofia, 1000 Sofia, 8 Kl. Ohridski Blvd, Bulgaria, meglena.laz@tu-sofia.bg

[3]Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Acad. G. Bonchev, Bl.8, Bulgaria, irgeorgiev@math.bas.bg, sggeorgiev@math.bas.bg

Abstract

Investment projects frequently undergo evaluation based on diverse criteria, including economic, social, environmental, and additional relevant factors [1, 2]. Each project is assigned points reflecting its performance in these criteria.

Due to limited budgets, not all qualified projects can be fully funded. This creates a challenging decision-making situation, where funding organizations need to select the best projects while maximizing both the number of approved projects and their overall quality. Additionally, practical funding decisions often result in partial rather than full financial support, requiring fractional resource allocation.

This paper proposes a multicriteria mixed-integer optimization model explicitly designed to allocate limited financial resources among competing investment projects by maximizing both the number of approved projects and their cumulative quality scores. The proposed approach allows partial funding, thereby enhancing flexibility and ensuring efficient utilization of available resources. The applicability and benefits of the proposed model are demonstrated through illustrative examples.

Key Words: *investment projects; multicriteria optimization; mixed-integer linear programming; resource allocation*

References

[1] Raeva I. (2019) Computation of risk in pricing of investment projects, AIP Conference Proceedings, Volume 2164, Article number 060016, DOI: 10.1063/1.5130818.

[2] Chakarov B. and Raeva I. (2022) Risk estimation using sensitivity analysis of an investment project, AIP Conference Proceedings, Volume 2522, Article number 060010, DOI: 10.1063/5.0101465.

MULTI-LEVEL THRESHOLDING IMAGE SEGMENTATION USING METAHEURISTIC ALGORITHM

Eda ÖZKUL[1], Muharrem KORKMAZ[1]

[1]Karadeniz Technical University, Faculty of Science, Department of Computer Sciences, Trabzon, Türkiye,
eda.ozkul.gs@gmail.com

muharremkrkmz11@gmail.com

Abstract

Image segmentation is the process of dividing an image into parts according to different features and making each part meaningful. Threshold-based segmentation is the most commonly used image segmentation method. These approaches are divided into two categories: bi-level and multi-level. The image is divided into two or more groups according to a threshold value by using threshold-based segmentation methods. In particular, they have been used extensively for grayscale images and are an advantage due to their low cost of computation. Multilevel thresholding, an extended version of bi-level thresholding, divides the image into many regions based on multiple thresholds. For this reason, multi-level thresholding is more appropriate for real-world problems. Otsu's method, Kapur entropy, Masi entropy, Renyi entropy, and Tsallis entropy are the most common thresholding methods. However, when these methods are used, the segmentation performance decreases as the threshold level increases, and the computational cost rises exponentially. In order to overcome this problem, the use of metaheuristic algorithms has become very widespread. In this study, an opposition-based metaheuristic algorithm is proposed. The proposed meta-heuristic algorithm is developed using the objective functions of the Otsu and Kapur thresholding methods and applied to the medical images. It is compared with well-known metaheuristics to evaluate the performance of the proposed algorithm. Experimental results show that the proposed algorithm is superior to the other algorithms in terms of MSE, PSNR, SSIM, and FSIM as performance metrics.

Key Words: *multilevel thresholding; Otsu; Kapur; metaheuristic algorithm; image segmentation.*

The role of banking heterogeneity in the relationship between financial development and entrepreneurship

Cristian BARRA[1], Christian D'ANIELLO[2],
(Christian D'ANIELLO[2])

[1]cbarra@unisa.it

[2]chdaniello@unisa.it

Abstract

Entrepreneurship plays a crucial role in economic progress as entrepreneurs, by investing and creating jobs, induce high economic growth in their country[1]. Since there are considerable start-up and running costs involved in starting up and running a business, the entrepreneurs seek large external loans at low interest rates. Financial intermediation can meet the financing needs of entrepreneurs because financial intermediaries can distribute funds at reasonable costs [2] and enable entrepreneurs to operate with minimal operating expenses boosting entrepreneurial activities that would otherwise fail due to too high operating costs. We employ a worldwide heterogeneous unbalanced sample consisting of 57 nations (developed and developing) from 2006 to 2021, considering information about two key types of financial intermediaries: commercial and cooperative. In our dataset the entrepreneurship is measured by the number of new business registrations per 1,000 persons aged 15-64; the financial development is measured by the number of commercial banks per 100,000 adults or cooperative banks and credit unions per 100,000 adults and we use as control variables the GDP, TRADE OPENNESS, EDUCATION, GOVERNMENT CONSUMPTION, TAXES POPULATION DENSITY. We propose both OLS and FE regression to explore the nexus under analysis and to assess whether the type of bank influences the relationship, in line with the view of Dutta and Meierrieks (2021) [3]. However one of the main limitations of OLS and FE approaches is endogeneity, resulting from omitted variables or reverse causality. To address such problems, the instrumental variables approach developed by Lewbel (2012) [4] is used. This method proves particularly useful when there is a lack of reliable external instruments ([5]; [3]) and also allows several instruments to be generated simultaneously. We include temporal dummies to control for omitted variables and cyclical factors, and also country dummies that allows us to capture sample heterogeneity. Our empirical study answer to two main research questions: (i) Does financial development support the entrepreneurship?, (ii) Is banking diversity necessary for financial development and entrepreneurship?. In this regard, the main contributions we offer to

the literature are the following: (i)Causal estimates of the effect of financial development on entrepreneurship using the IV-H approach, devised by Lewbel (2012), (ii)Considering banking heterogeneity. Based on three separate estimators (OLS, FE and IV-H), the empirical evidence confirms the relevance of both financial intermediaries in supporting entrepreneurship, with commercial banks having a higher intensity than credit unions and cooperatives. In this sense, despite the differences in the missions of these two types of financial institutions, regulators and policymakers should promote procedures that encourage collaboration between the two types of institutions.

Key Words: *Entrepreneurship; Financial Development; Bank Diversity; International Heterogenous Sample; IV-H*

References

[1] ECB's 2018 Euro Area Corporate Finance Survey

[2] Levine, R. (1997). Financial development and economic growth: Views and agenda. *Journal of Economic Literature*, 35, 688-726.

[3] Dutta, N, & Meierrieks, D. (2021). Financial development and entrepreneurship. *International Review of Economics and Finance*. 73, 114-126.

[4] Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30, 67-80.

[5] Arcand, J. L., Berkes, E., & Panizza, U. (2015). Too much finance? *Journal of Economic Growth*, 20, 105-148.

Stochastic properties of sums of dependent random variables (risks)

Prof. Dr. Jorge NAVARRO
University of Murcia, SPAIN

Abstract:

We will see some recent results on stochastic ordering properties for sums of dependent random variables. These results can be applied to risks and to lifetimes in survival analysis or reliability theory. They are based on copula representations for this kind of sums (C-convolutions) and on distortion representations. We will get results for the usual stochastic order (also called first-order stochastic dominance), the increasing concave order (second-order stochastic dominance), the increasing convex order, the hazard rate order and the convex order. Some illustrative examples will be considered.

Acknowledgements:

Partially supported by Ministerio de Ciencia e Innovación of Spain in the project PID2022-137396NB-I00, funded by MICIU/AEI/10.13039/501100011033 and by 'ERDF A way of making Europe'.

References:

- Navarro J., Sarabia J.M. (2022). Copula representations for the sums of dependent risks: models and comparisons. Probability in the Engineering and Informational Sciences 36, 320–340. Doi: 10.1017/S0269964820000649.
- Navarro J., Pellerey F., Mulero J. (2022). On sums of dependent random lifetimes under the time-transformed exponential model. Test 31, 879–900. Doi: 10.1007/s11749-022-00805-2
- Navarro J., Zapata J.M. (2025). Stochastic dominance of sums of risks under dependence conditions. Submitted. ArXiv: 2503.05348v1.

A NOTE ON A CLASS OF ESTIMATES FOR THE REGRESSION FUNCTION

Mounir ARFI

Sultan Qaboos University, College of Science, Department of Statistics, P.O. Box 50, Al Khoodh, Postal Code No 123, Muscat, Sultanate of Oman, arfi@squ.edu.om

Abstract

The nonparametric techniques used in the kernel estimation of the regression function have been widely investigated by many researchers since Watson-Nadaraya (1964) who applied the regression function to some climatological time series data by means of a kernel function.

Since then, a large number of authors were interested in the kernel estimation of the regression function.

In many stochastic models, the assumption that the random variables are independent is not plausible and the extent to the dependence case seems to be more convenient. One of those dependence structures is the rho-tilde mixing condition.

In our work, we obtain some results for a class of kernel estimates of the regression function when the observed process is rho-tilde mixing and when the data belong to a sequence of compact sets that increases to the whole space as *n* grows to infinity.

Key Words: *regression function; kernel estimation; rho-tilde mixing condition; increasing sequence of compact sets as n grown to infinity; almost sure convergence*

References

[1] Bradley, R.C., Equivalent mixing conditions of random fields, Technical Report 336, Center of Stochastic Processes. University of North Carolina, Chapel Hill(1990).

[2] Guang-Hui, C., Strong law of large numbers for $\tilde{\rho} - \text{mixing}$ sequences with different distributions, Discrete Dynamics in Nature and Society, (2006), 1-17.

[3] Meng-Hu, Z., C., Strong law of large numbers for arrays of raw-wise $\tilde{\rho} - \text{mixing}$ random variables, Discrete Dynamics in Nature and Society, (2007), 1-17.

[4] Nadaraya, E. A., On estimating regression, Theory of Probability and its Applications, 9, (1694), 141-142.

[5] Shixin, G., Almost sure convergence for $\tilde{\rho} - \text{mixing}$ random variable sequences, Statistics & Probability Letters, 67, (2004), 289-298.

[6] Watson, G.S., Smooth regression analysis, Sankhya A, 26, (1964), 359-372.

Use of linear dynamic models for time series prediction

Luboš MAREK[1], Richard HINDLS[2], Stanislava HRONOVÁ[3]

[1]Prague University of Economics and Business, Czech Republic, marek@vse.cz

[2]Prague University of Economics and Business, Czech Republic, hindls@vse.cz

[3]Prague University of Economics and Business, Czech Republic, hronova@vse.cz

Abstract

The aim of the paper is to show how time series dependencies and time lags of these dependencies can be used to construct forecasts. As a suitable class of models, we have chosen linear dynamic models, which allow us to describe the dependency between one time series in the form of an explanatory variable (output series) and one or more time series that appear in the form of explanatory variables (input series). We are interested in models that are made up of economic time series. It is necessary to use a tool that measures not only the intensity but also the direction of the dependency. Therefore, we use the cross-correlation coefficient. Specifically, we will use models with a transfer function to construct the model.

In economics, it is common to know the values of some time series immediately. For example, the daily exchange rate of CZK (Czech crown) against the dollar and the Euro is known essentially on the second calendar day. Similarly, the average monthly exchange rate is known immediately after the end of the calendar month. In contrast, for example, the Czech Statistical Office publishes import and export values with a two to three months delay. If it were possible to model exports (imports) depending on the CZK/USD or CZK/EUR exchange rate, it would be possible to estimate the values of imports and exports in advance and use them to estimate the gross domestic product more quickly. Creating a model with a time lag allows the creation of forecasts when the real observed values of the explanatory variable (input series) are inserted into the model. Such a procedure will significantly shorten the estimation (publication) time of these important economic series, but will also make the forecasts of other macroeconomic aggregates more accurate. There are many similar examples of the interdependence of economic time series, including lags, as we will show in our contribution.

Key Words: *time series; linear dynamic model; transfer function models; CCF;*

References

[1] Ferreira Lopes CE, Cross NJG (2017) New Insights into Time Series Analyses. A&A 604, A121.
[2] Fox J (2016) Applied regression analysis and generalized linear models. Sage Publications, London.
[3] Greene WH (2012) Econometric Analysis, 7th ed. Pearson Education. Prentice Hall.
[4] Pankratz, A (1991) Forecasting with dynamic regression models. John Wiley & Sons inc., New York.
[5] Wei, WWS (1990) Time series analysis – Univariate and multivariate methods. Redwood City, California, Addison-Wesley Publishing Company.
[6] Woodward WA, Gray HL, Elliott AC (2012) Applied Time Series Analysis. CRC Press.

A FUZZY C-MEANS IMAGE SEGMENTATION ALGORITHM USING METAHEURISTICS

Özge TEZEL[1], Onur OKUYUCU[1]

[1]Karadeniz Technical University, Faculty of Science, Department of Computer Sciences, Trabzon, Türkiye, ozge_tzl@hotmail.com

onurokuyucu1@gmail.com

Abstract

Image segmentation, a significant step in image processing, is defined as dividing an image into meaningful parts or regions according to specific features. Nowadays, it is widely used in medical images, satellite images, pattern recognition, image analysis, and security systems. The fuzzy c-means (FCM) algorithm is often used for image segmentation because of its simplicity and efficiency. However, the FCM algorithm has some drawbacks. Especially in problems with large and complex data sets, the computational cost increases due to the huge search space and its sensitivity to initial values. However, the algorithm is sensitive to noisy data, which limits its effectiveness and leads to unsatisfactory results. In this study, a hybrid metaheuristic algorithm is proposed to deal with the disadvantages of the FCM. The proposed metaheuristic algorithm is based on FCM's objective function. The proposed algorithm is applied to benchmark and medical images used in image processing. The proposed method is compared with well-known metaheuristic algorithms using cluster validity indices to demonstrate the efficiency of the algorithm. The proposed hybrid approach provides better results in image segmentation as it is less sensitive to initial values and noise. Moreover, the obtained results reveal the superiority of the proposed method over the compared methods.

Key Words: *image segmentation; fuzzy c-means; cluster validity index; metaheuristics.*

PREDICTING MISSING DATA USING INACTIVITY TIMES OF COMPONENTS UPON SYSTEM FAILURE

Anna DEMBIŃSKA[1]

[1]Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warsaw, Poland, e-mail: anna.dembinska@pw.edu.pl

Abstract

Inactivity time of an object (an item or a system), also known as passed lifetime, is an important concept studied in the reliability theory. It describes the time that elapsed from the failure of the object to time *t*, given that it failed before time *t*. This concept is dual to another useful quantity in survival analysis, namely to residual lifetime of an object. The latter represents the time that passed from the moment *t* till the failure of the object, given that the object was working at the moment *t*.

During the talk, motivated by missing data problems, I will generalize the concept of inactivity time of an item to inactivity times of components upon system failure. More precisely, for a system comprising of *n* components, I will define inactivity times of components upon system failure as times that elapsed from failures of non-surviving components to the breakdown of the whole system. I will show that these inactivity times do not depend on the structure of the system provided that the joint distribution of component lifetimes is exchangeable and absolutely continuous. I will also present a handy formula expressing the joint survival function of the inactivity times of interest. Next, I will use a copula to model the dependence structure among components and will express the formula for survival function of inactivity times in terms of this copula. Furthermore, I will demonstrate numerical results for some special cases with specific copulas and margins, and will derive limiting properties as the failure time of the system tends to infinity. Finally, I will give an example of application of the presented results to predicting missing medical data.

The talk will be based on the joint work with Katherine Davis from McMaster University [1].

Key Words: *inactivity time; missing data; reliability theory; exchangeability; copula*

References

[1] Davies K, Dembińska A (2025) Inactivity Times of Dependent Components Upon System Failure with Application to Missing Data Problems, *manuscript.*

Using HMD Methods Protocol to Estimate Turkish Population at the Advanced Ages

Özer BAKAR[1], Murat BÜYÜKYAZICI[2], İsmet KOÇ[3]

[1]Hacettepe University, Department of Actuarial Sciences, ozerbakar@hacettepe.edu.tr
[2]Hacettepe University, Department of Actuarial Sciences, muratby@hacettepe.edu.tr
[3]Hacettepe University, Institute of Population Studies, iskoc@hacettepe.edu.tr

Abstract

The Human Mortality Database (HMD) [1] is the world's leading mortality data source in developed countries. HMD provides researchers, policy analysts and whoever interested in human longevity, easy access to high quality, harmonized and detailed data. Also, database consists input data as vital statistics, censuses and official population estimates along with the step by step calculations called as methods protocol which makes the database transparent and trustworthy in the scientific area. Turkish Statistical Institute (TurkStat) [2] provides mortality data upto age 98 for deaths and upto age 90 for population counts (for thoes within de 5 year age group). These age limits are much lower than the upper age limit: 110, which is the HMD standart for each nation in the database. In this study, the statistical methods in the HMD protocol is applied to the Turkish mortality and population data between the years 2009 and 2020 and the number of deaths and the population counts at the advanced ages are estimated.

Keywords: *Mortality, Population Counts, Human Mortality Database, Advanced Age.*

Acknowledgements:

This study has been supported by the TUBİTAK 1001 with the grant number 123R017 and by the Scientific Research Projects Coordination Unit of Hacettepe University under grant number SHD-2024-20912.

References

[1] HMD, (2024). Human Mortality Database. Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA) and French Institute for Demographic Studies (France). Available at www.mortality.org.

[2] Turkstat, (2024). Turkish Statistical Institute.

ON A SPLITTING METHOD FOR SOLVING OF THE NONLINEAR SCHROEDINGER EQUATION AND ITS GENERALIZATION TO THE MANAKOV SYSTEM

Michail TODOROV[1], Meglena LAZAROVA[2]

[1]Institute of Mathematics and Informatics, BAS, 1113 Sofia, Akad. G. Bonchev bl.8, Bulgaria, mdtodorov@math.bas.bg

[2]Faculty of Applied Mathematics and Informatics, Technical University of Sofia, 1000 Sofia, 8 Kl. Ohridski Blvd, Bulgaria, meglena.laz@tu-sofia.bg

Abstract

In this paper we aim to demonstrate that the splitting by physical factors [1] is applicable and can be efficient for solving of the nonlinear Schrodinger equation [2]. Both the linear and nonlinear parts are treated by the Runge-Kutta method, the nonlinear term, however, is linearized by the so-called inner iteration. By adding of a coordinate splitting of the spatial coordinates the method can be expanded and relatively easily applied for (2+1)d Schroedinger equation as well as to the Manakov system. Then the procedure should be applied for each equation in the system. The conducted numerical simultions and their results are reliable and give good predictions for the material quantities and dynamics of the ligh. They give very good comparison with the previous papers of the authors got in another methods [3,4].

Key Words: *nonlinear Schroedinger equation; Manakov system; splitting by physical factors; numerical simulations*

References

[1] Marchuk G. Mathematical Models in Environmental Problems, Studies in Mathematics and Its Applications, ed J L Lions et al. Vol.16 (Amsterdam North-Holland, 1986) pp 82-9. [2] Todorov M. On a Method for Solving of Multidimensional Equations of Mathematical Physics, ed M Belhaq, CSNDD 2016, Marrakech, Morocco, EDP Sciences, France MATEC Web of Conferences, vol.83, paper 05012, 3p., doi:10.1051/matecconf/20168305012.

[3] Gerdjikov VS, Todorov MD, and Kyuldjiev AV (2016) Mathematics and Computers in Simulation 121, 166-178, doi: 10.1016/j.matcom.2015.10.004

[4] Gerdjikov VS and Todorov MD (2019) Applied Numerical Mathematics 141, 62-80, doi: https://doi.org/10.1016/j.apnum.2018.05.015.

Development of a Deep Learning-Based Hybrid Model for Diagnosing Lung Cancer Using CT Scan Images

Yasemin GÜNTER[1], Ayşenur ÖZTÜRK[2], Fatma NOYAN TEKELİ[3]

[1] Hamidiye, Anadolu Cd. No:40, 34408 Kağıthane/İstanbul, yasemin.gunter@atlas.edu.tr

[2] Hamidiye, Anadolu Cd. No:40, 34408 Kağıthane/İstanbul, aysenur.ozturk@atlas.edu.tr

[3] YTÜ-Davutpaşa Kampüsü, 34220 Esenler/İstanbul, fnoyan@yildiz.edu.tr

Abstract

Lung cancer is one of the most common and life-threatening types of cancer worldwide. In 2022, approximately 2.5 million new cases of lung cancer were diagnosed globally, accounting for 12.4% of all new cancer cases. In the same year, approximately 31,000 people in Turkey were diagnosed with lung cancer, and nearly 25,000 lost their lives due to this disease. These statistics highlight that lung cancer remains a significant health concern both globally and in Turkey. Early diagnosis and effective treatment methods are critical in managing the disease. Computed Tomography (CT) scans stand out as one of the most effective methods for diagnosing lung cancer in the healthcare sector. However, identifying and accurately interpreting tumors in CT scans remains a challenge for doctors and radiologists. At this point, Artificial Intelligence (AI) technologies come into play, facilitating the diagnosis of lung cancer through CT scans. Studies in this field aim to accurately detect abnormalities in targeted images. This study focuses on the early and accurate diagnosis of lung cancer from CT scans by leveraging Deep Learning (DL) strategies to enable rapid and precise image analysis. A hybrid model along with four deep learning models is proposed for the early detection of lung cancer in this research. To further enhance the performance of the proposed models, the hyperparameters will be automatically tuned using Grey Wolf Optimization (GWO). The models, namely GWO-VGG16, GWO-DenseNet, GWO-DenseNet+LSTM, GWO-InceptionV3, and a hybrid approach called GWO-VGG16+InceptionV3, will be applied to CT images. [3] Deep learning and Grey Wolf Optimization (GWO) offer remarkable advantages in the field of image classification. These advantages include high accuracy rates, the ability to automatically extract features, the capacity to work with large datasets, transfer learning capabilities, robustness against variable conditions, the discovery of optimal solutions, automatic hyperparameter tuning, and continuous progress achieved through ongoing research and development. By leveraging the capabilities of deep learning and GWO in image processing, the study aims to develop a faster and more effective lung cancer diagnosis model. The proposed four models and the hybrid approach will be compared using various performance metrics such as precision, accuracy, specificity, sensitivity, and AUC. As a result of these comparisons, a deep learning model focused on the early detection and classification of lung cancer using CT scan images will be proposed.

Key Words: *CT, lung cancer, Deep Learning, VGG-16, InceptionV3.*

References

[1] Ozcelik N, Kıvrak M, Kotan A, Selimoğlu İ (2024) Lung cancer detection based on computed tomography image using convolutional neural networks. Technology and Health Care. 2024;32(3):1795-1805. doi:10.3233/THC-230810.

[2] Tárnoki ÁD, Tárnoki DL, Dąbrowska M, Knetki-Wróblewska M, Frille A, Stubbs H, et al.

(2024) New developments in the imaging of lung cancer. Breathe (Sheff). 2024;20(1):230176. doi:10.1183/20734735.0176-2023

[3] Kumaran SY, Jeya JJ, RM T, et al. (2024) Explainable lung cancer classification with ensemble transfer learning of VGG16, Resnet50 and InceptionV3 using grad-cam. BMC Med Imaging. 2024;24(176):1-10. doi.org/10.1186/s12880-024-01345-x

[4] Nayak T, Gokulkrishnan N, Chadaga K, Sampathila N, Mayrose H, KS S (2024) Automated histopathological detection and classification of lung cancer with an image pre-processing pipeline and spatial attention with deep neural networks. Cogent Engineering. 2024;11(1):2357182. doi.org/10.1080/23311916.2024.2357182

[5] Shafi I, Din S, Khan A, De La Torre Díez I, Palí Casanova RdJ, Tutusaus Pifarre K, et al. (2022) An effective method for lung cancer diagnosis from CT scan using deep learning-based support vector network. Cancers (Basel). 2022;14(21):5457. doi:10.3390/cancers14215457.

[6] Musthafa MM, Manimozhi I, Mahesh TR, et al. (2024) Optimizing double-layered convolutional neural networks for efficient lung cancer classification through hyperparameter optimization and advanced image pre-processing techniques. BMC Med Inform Decis Mak. 2024;24(142):1-10. doi.org/10.1186/s12911-024-02553-9

AIRPORT CUSTOMER SEGMENT RANKING: HYBRID TOPSIS APPROACH

Beyza AHLATÇIOĞLU ÖZKÖK[1], Serhan KARADENİZ[2], Sezin ÖZTÜRK[3]

[1] Department of Business Administration, Faculty of Economics and Administrative Sciences, Yildiz Technical University, Davutpasa, Istanbul, Türkiye, e-mail: bahlat@yildiz.edu.tr

[2]Aviation Management Department, School of Civil Aviation, Beykoz University, Beykoz, Istanbul, Türkiye, e-mail: serhankaradeniz@beykoz.edu.tr

[3]Department of Business Administration, Faculty of Economics and Administrative Sciences, Yildiz Technical University, Davutpasa, Istanbul, Türkiye, e-mail: sezin.ozturk@yildiz.edu.tr

Abstract

Airports are one of the essential infrastructures that significantly contribute to contemporary society. The assessment and ranking of the performance and efficiency of these decision-making units, which have a significant societal impact, as well as the identification of their strengths and weaknesses, have long been the subject of research [1]. The aim of the current research is to measure and rank the efficiency of airports by incorporating service quality and sentiment analysis data alongside operational data. As the sample, airports in Europe with a passenger traffic of 20 million or more in 2022 were selected. Due to some data limitations, the final sample consisted of 20 airports. For operational data, airports have been evaluated based on total passenger numbers, aircraft movements, cargo tonnage, and CO2 emissions for the years 2021-2023, while passenger reviews from the same period have been utilized for service quality assessment and sentiment analysis. Customer reviews were categorized according to passenger segments on SKYTRAX, and as a result, separate rankings were created for the All, Solo, Business, Couple, and Family segments. Airports with lower mean rankings and low standard deviations, such as Leonardo da Vinci–Fiumicino Airport and Zurich Airport, exhibit strong and consistent performance across all passenger segments. These airports can be considered benchmarks of balanced service quality. Airports with low mean rankings but high standard deviations, such as Istanbul Airport, demonstrate strong performance in certain segments while needing targeted improvements in others. This highlights the importance of tailoring services to meet the needs of diverse traveler groups. Additionally, the analyses reveal which airports have lower rankings in certain categories and higher rankings in others. Airports like Lisbon Airport display high mean rankings with low standard deviations, indicating consistent underperformance across segments. Such cases sug- gest systemic issues requiring comprehensive improvements. Airports such as Adolfo Suarez Madrid–Barajas Airport and Oslo Airport, Gardermoen, with high standard deviations, experience significant ranking variability. These findings suggest that service quality varies significantly depending on the passenger type, necessitating a segmented approach to service enhancement. As a result of the study, a new method was utilized for airport sector efficiency analysis, incorporating service quality and sentiment analysis scores alongside operational data. This approach allows airports to not only determine their overall position but also identify areas for improvement within specific customer segments.

Key Words: *Aviation Industry; Airports; Sentiment Analysis; MCDM; TOPSIS*

References

[1] See, K. F., Ülkü, T., Forsyth, P., & Niemeier, H. M. (2023). Twenty years of airport efficiency and productivity studies: A machine learning bibliometric analysis. Research in Transportation Business & Management, 46, 100771.

THE INTEGRATION OF ARTIFICIAL INTELLIGENCE TOOLS IN TEACHING STATISTICS

Laura ASANDULUI[1], Christiana SANDU[2]

[1]e-mail: asand@uaic.ro

[2]Author's contact address and e-mail: christiana.balan@uaic.ro

Abstract

In the artificial intelligence and data-driven era, teaching statistics has become increasingly challenging. Classical statistical methods, which are commonly used in statistics education, rely on probability theory and require specific assumptions about data distributions. Artificial intelligence tools focus on large datasets that are hypotheses free.

The purpose of teaching traditional statistics is to develop statistical thinking and to provide an understanding of statistical methods. Traditional statistics focus on statistical concepts, problem-solving skills, data collection methods (e.g. sampling design), and the appropriate statistical methods. Instructors can help students understand statistical methods and reasoning, as well as interpret the results.

The challenge for instructors is to integrate artificial intelligence tools into statistics education while leveraging the advantages of both traditional statistical methods and artificial intelligence approaches. Artificial intelligence tools can handle larger, more complex datasets and reveal elaborate patterns. Teaching fundamental statistics fosters rigorous analytical skills, while artificial intelligence tools provide new possibilities for analyzing complex data.

Key Words: *teaching statistics; artificial intelligence tools; classical statistical methods*

References

[1] Baker RS, Inventado PS (2014) Educational data mining and learning analytics in J. A. Larusson and B. White (Eds.), Learning Analytics: From Research to Practice 61–75.

[2] Holmes W, Bialik, M, and Fadel C (2019), Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.

[3] Luis F Rivera-Galicia (2018) The Challenges of Teaching Statistics to Undergraduate Business and Economics Students in Spain, ICOTS10

[4] Schwarz J (2025) The use of generative AI in statistical data analysis and its impact on teaching statistics at universities of applied sciences, Teach. Stat. 1–11

[5] Zavez K, and Harel O (2025) Teaching Statistical Concepts Using Computing Tools: A Review of the Literature. Journal of Statistics and Data Science Education, 1–12. https://doi.org/10.1080/26939169.2024.2445541

[6] Wan Saiful 'Azzam Bin Wan Ismail, Nur Amalina Bt Awang, Nooradzlina Bt Mohd Pauzi (2024) An Overview of The Use of Artificial Intelligence (AI) Tools in Higher Education, International Journal on e-Learning and Higher Education Volume 19, Number 3.

InVEnT 2025

The 5th International Conference on Artificial Intelligence, Smart Technologies and Engineering Applications

The 5th International Conference on

Artificial Intelligence, Smart Technologies and Engineering Applications

Conference Theme: AI Beyond the Comfortable, Secure and Sustainable Society

25 – 27 April 2025 Bologna – ITALY

Dear Colleagues,

The 5th International Conference on Artificial Intelligence, Smart Technologies and Engineering Applications (INVENT-2025) aims to provide an excellent international forum for sharing the research and developments, and their results in the constantly evolving fields of Artificial Intelligence, intelligent systems and application of Intelligent Systems in various engineering fields. The first four episodes of the INVENT conference focused on smart energy systems and applications. However, the conference organizing committee decided to broaden the scope of the conference and in addition to smart energy systems, cover the new smart technology trends such as AI, machine learning and engineering applications in order to bring together larger number of distinguished scientists and enhance the degree of interaction.

The INVENT-2025 conference will cover Artificial Intelligence, Machine Learning, Pattern Processing, Smart Energy Systems, Automation and Optimization, Robotics, IoT, Sensors, Telecommunication Networks and Systems, Biomedical Systems and Engineering Innovations, Data Science and Analytics, Artificial Intelligence and Cyber Security, Smart Cities, Human-Computer Interaction , User Experience Design, Block-Chain Applications in Industry, Sustainable Production with Smart Technologies, VR, AR, ER, Data Processing with Sensors, Evolutionary Calculations (Genetic Algorithms, Evolution Strategies, Particle Swarm Optimization etc.), Processing of Electromagnetic and Acoustic Waves, GAI in Education.

The conference will provide a platform where the latest developments, research results, projects and industrial experiences in the above-mentioned fields will be discussed by researchers working in both academia and industry. Authors are invited to contribute to the conference by submitting articles showing their work in relevant fields.

Submitted papers will be peer-reviewed by the technical committee members. Accepted and registered conference papers will be published in the conference e-proceedings in the MDPI Computer Sciences & Mathematics Forum journal (ISSN 2813-0324) if requested by the author. for more information please visit publication page

We kindly wait for your attendance to our conference to be held on 25 – 27 April 2025 in Bologna, ITALY

All informations are available in conference web site. For more information please do not hesitate to contact us. info@inventconference.com

Prof. Dr. Hasan AMCA
International Conference Chair

International Conference Chair
Prof. Dr. Hasan Amca

Conference Co-Chairs
Prof. Dr. Mustafa Uyguroglu (Robotics, Eastern Mediterranean University),
Prof. Dr. Hasan Demirel (Image Processing, Eastern Mediterranean University),
Prof. Dr. Uğur Atikol (Smart Energy, Eastern Mediterranean University),
Prof. Dr. Qasim Zeeshan (ML, Eastern Mediterranean University),
Prof. Dr. Orhan Korhan (Industry 5.0, Eastern Mediterranean University),
Prof. Dr. Hüseyin Ademgil (Optical Comm, European University of Lefke)
Prof. Dr. Bülent Bilgehan (Smart Communications, Near East University),

Technical Program Committee
Prof. Dr. Murat Fahrioğlu (Middle East Technical University),
Assoc. Prof. Dr. M. Atif SHAHZAD (Dept. of Industrial Eng., King Abdulaziz University),
Assoc. Prof. Dr. Alper Doğanalp (Eastern Mediterranean University),
Assoc. Prof. Dr. Samet Biricik (Near East University),
Assist. Prof. Dr. Mustafa Mulla (Cyprus International University),
Dr. Suat Yeldener (Freelance Entrepreneur),
Dr. Zeki Murat Çınar (Groupe Atlantic, İzmir, Türkiye),
Asst. Prof. Dr. Oğuzhan Kırılmaz (Eastern Mediterranean University)

Smart Classrooms of the Future: AI-Driven Support for Teachers and Students †

Laura Pagani [1]*, Paolo Gemelli [2]

[1] Università degli Studi di Genova, Dipartimento di Architettura e Design, ITALY, laura.pagani@edu.unige.it

[2] Università degli Studi di Genova, Dipartimento di Architettura e Design, ITALY,paolo.gemelli@edu.unige.it

* Correspondence: laura.pagani@edu.unige.it Tel. +39 3471091637

Abstract: Artificial Intelligence (AI) is revolutionizing education by enhancing personalized learning and student well-being. This study presents an **AI-Driven Classroom Council**, a multi-agent system designed to support teaching, foster collaboration, and monitor students' emotional states. Specialized AI agents—virtual teachers, tutors, psychologists, and SEN advisors—work together to create an inclusive learning environment. Case studies highlight AI-assisted **STEM support** and **concept mapping for SEN students**. Ethical concerns are addressed through **expert supervision**. Findings show **reduced academic stress, improved engagement, and personalized learning**. This research explores AI's potential to transform digital education while ensuring fairness, privacy, and scalability.

Keywords: GAI assisted education, Personalized Learning, Student Well Being

1. Introduction

The integration of Artificial Intelligence (AI) in education is transforming traditional learning environments, offering unprecedented opportunities to enhance **personalised learning, student well-being, and teaching efficiency** [5]. AI-driven solutions are reshaping the educational landscape by providing adaptive learning experiences, real-time student monitoring, and personalised interventions. With AI's growing role in learning, educators and policymakers seek innovative methods to implement AI responsibly and ethically within classrooms [4].

This paper introduces the **AI-Driven Classroom Council**, a multi-agent AI-based system designed to support teachers, enhance student engagement, and provide **individualised learning pathways**. The system includes **AI-powered virtual educators, tutors, psychologists, and special education advisors (SEN agents)** who collaborate to ensure students receive appropriate, tailored support. This research presents the system's architecture, explores its potential impact through case studies, and discusses the ethical and practical implications of deploying AI in education. The aim is to demonstrate how AI enhances **learning efficiency, emotional support mechanisms, and classroom management strategies** while ensuring that human educators remain central in the learning process.

.

2. AI-Driven Classroom Council Architecture

The AI-Driven Classroom Council is a multi-agent system where specialised AI entities work together to optimise classroom management, teaching methodologies, and student assistance [6]. These agents interact dynamically to assess student needs, adapt educational content, and monitor cognitive engagement. The system is structured to allow seamless interaction between AI components and human educators, reinforcing the teacher's role in the classroom while automating labour-intensive tasks.

2.1 System Components and Interactions

The system comprises the following key agents:

- **AI Principal Agent**: Oversees system functionality, ensuring compliance with educational standards and optimising pedagogical methodologies.
- **AI Subject Teachers**: Provide adaptive instruction, tailored lesson materials, and real-time feedback based on student performance analytics [8].
- **AI Tutor Agent**: Offers **homework assistance, concept clarification, and individualised study plans** based on detected student learning gaps.
- **AI SEN Support Agent**: Customises learning resources for students with **Special Educational Needs (SEN)**, ensuring accessibility and inclusivity [3].
- **AI Psychological Support Agent**: Uses **sentiment analysis, linguistic markers, and biometric data** to detect stress or disengagement, offering psychological interventions [1].
- **AI Cognitive Monitoring Agent**: Tracks cognitive load, attention levels, and engagement using EEG, eye-tracking, and biometric feedback to assess student focus and prevent burnout [2].
- **AI Coordinator Agent**: Manages data exchange between agents, ensuring that interventions are timely and optimally aligned with student progress.

Each agent continuously communicates with teachers and administrators, ensuring that AI recommendations align with human expertise and pedagogical best practices.

2.2 Enhanced Learning Cycle

The AI-driven learning cycle follows a continuous feedback loop:

1. **Data Collection –** AI monitors student engagement, cognitive state, and academic performance.
2. **Processing & Analysis –** AI aggregates and interprets student data, detecting potential struggles or areas of improvement.
3. **Intervention & Adaptation –** AI suggests tailored content modifications and real-time interventions.
4. **Teacher Oversight & Decision-Making –** Educators review AI suggestions and implement personalised learning strategies.
5. **Evaluation & Optimization –** AI tracks the success of interventions, refining future recommendations.

This framework ensures a balance between AI-driven insights and human expertise, maintaining teacher authority while maximising student success.

3. Teacher-AI Collaboration: Benefits for Educators and Students

One of the primary concerns in the adoption of AI in education is its perceived role in replacing teachers. However, the AI-Driven Classroom Council is designed to function as an enhancement tool rather than a replacement, focusing on augmenting teachers' capabilities while ensuring that human educators remain central in the learning process.

The collaboration between AI and teachers fosters a dynamic educational environment where both parties complement each other's strengths. AI provides data-driven insights, automates repetitive tasks, and offers adaptive learning recommendations, but the human element ensures creativity, emotional intelligence, and mentorship, which are irreplaceable aspects of teaching.

The AI-Driven Classroom Council creates a symbiotic relationship where teachers can use AI-generated analytics to tailor their instruction more precisely. This hybrid approach reduces teacher burnout by minimizing administrative burdens while enabling real-time identification of struggling students. AI also enhances teachers' ability to implement differentiated instruction, ensuring that each student progresses at their own pace.

Table 1: Comparison Between Traditional and AI-Assisted Teaching

Aspect	Traditional Teaching	AI-Assisted Teaching
Lesson Planning	Manual preparation	AI suggests personalized adjustments
Student Monitoring	Based on teacher observation	AI tracks cognitive load & emotions
Administrative Tasks	Time-consuming	Automated grading & attendance
Early Interventions	Delayed detection	AI flags potential learning issues in real-time
Feedback Mechanism	Periodic evaluations	Continuous AI-driven feedback loops
Engagement Strategy	Static lessons	Gamification & interactive learning

AI and teachers working together enhance student engagement, streamline curriculum adjustments, and foster personalized learning environments that benefit both students and educators alike.

4. Ethical Considerations & Data Privacy

With AI's expansion in education, ethical concerns and regulatory compliance are critical factors. To address privacy risks and biases, the AI-Driven Classroom Council ensures:

- **Transparency and Explainability**: AI-driven decisions are clearly interpretable and communicated to students and educators.
- **Bias Detection & Fairness Audits**: AI models undergo regular audits to eliminate discriminatory biases related to learning abilities, disabilities, or socio-economic factors.
- **Privacy & Data Protection**: Student data is anonymized, encrypted, and stored securely in compliance with GDPR and FERPA regulations [4].
- **Human-Centric AI Design**: AI augments teachers' roles rather than replacing them, fostering an AI-assisted but human-led classroom environment.

5. Future Directions and Implementation Challenges

As AI-based education expands, several challenges need to be addressed to enhance its real-world applicability and ensure equity in learning outcomes. The following areas require further research and development:

- **Longitudinal Studies:** To assess AI's long-term impact on student success rates, retention, and emotional well-being, large-scale studies must be conducted in diverse educational settings.
- **Expanding Adaptive Learning Capabilities:** Future work should improve AI's ability to cater to different cognitive styles, learning disabilities, and multilingual environments.
- **Teacher Training Programs:** Effective AI implementation requires extensive training for educators to seamlessly integrate AI tools into their teaching methodologies while maintaining authority over pedagogical decisions.
- **Regulatory and Ethical Considerations:** Policies need to be developed to guide responsible AI use in education, ensuring data privacy, algorithmic fairness, and student rights protection.
- **Scalability & Infrastructure Development**: Future AI-driven classroom models should be designed to scale across various education systems worldwide while considering technological limitations in lower-resource settings.

Additional Considerations:

- **Implementation and Scalability**: While AI-driven solutions promise to transform education, their implementation in traditional school settings remains a challenge. Many schools lack the technological infrastructure and financial resources to support such advanced AI systems. Furthermore, integration with existing curricula and teaching methodologies may require significant restructuring.
- **Dependence on AI**: Despite emphasizing AI as a supportive tool rather than a replacement, there remains a risk that over-reliance on AI could diminish the role of human educators. The irreplaceable aspects of teaching—such as emotional intelligence, mentorship, and creative adaptability—must remain central in AI-assisted learning environments.
- **Privacy and Security of Data**: The collection and analysis of biometric data (such as EEG tracking, eye movement monitoring, and sentiment analysis) introduce potential privacy risks. Ensuring compliance with GDPR and other privacy regulations is crucial to protect students from potential misuse of their sensitive data.

The collaboration between teachers and AI is not about replacement but empowerment—giving educators the tools they need to provide effective, personalized, and engaging learning experiences for students in the digital age.

Author Contributions: Conceptualization L.P. and P. G.; methodology, L.P. and P. G.; formal analysis, L.P. and P. G; investigation, L.P. and P. G; resources, L.P. and P. G.; data curation, L.P. and P. G.; writing—original draft preparation, L.P. and P. G.; writing—review and editing, L.P. and P. G.; visualization, L.P. and P. G.; supervision, L.P. and P. G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

SEN Special Educational Needs

References

1. Baker, R. S., D'Mello, S. K., Rodrigo, M. M., & Graesser, A. C. (2018). Better to be frustrated than bored: The incidence and persistence of affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223-241.
2. D'Mello, S. K., Dieterle, E., & Duckworth, A. (2020). Advanced, analytic, automated (AAA) measurement of engagement during learning. *Educational Psychologist, 55*(1), 1-17.
3. Feng, M., & Heffernan, N. (2020). The impact of adaptive learning technology on student engagement and performance. *Computers & Education, 153*, 103896.
4. Holmes, W., Bialik, M., & Fadel, C. (2021). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Routledge.
5. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2018). *Artificial Intelligence and Future Education: Promises and Challenges*. UCL Institute of Education Press.
6. Nguyen, A., Gardner, M. K., & Sheridan, K. (2020). AI-powered adaptive learning in K-12 education. *Journal of Learning Analytics, 7*(2), 112-130.
7. Van den Broek, G. S., Takashima, Y., & Yoshikawa, H. (2021). AI-assisted teacher professional development. *Technology, Pedagogy and Education, 30*(1), 22-39.
8. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on AI in education. *International Journal of Educational Technology in Higher Education, 16*(1), 1-27.

Hybrid Waste Management System: Sensor and Vision-based Data Collection and Artificial Intelligence-based Classification

Bedirhan Gürsu, Berkant Özkan, Eren Okumuş, Ömer Faruk Gülhan, Özden Niyaz

1st Bedirhan Gürsu
Department of Electronics and Communication Engineering
Yıldız Technical University
İstanbul, Türkiye
bedirhan.gursu@std.yildiz.edu.tr

2nd Berkant Özkan
Department of Electronics and Communication Engineering
Yıldız Technical University
İstanbul, Türkiye
berkant.ozkan@std.yildiz.edu.tr

3rd Eren Okumuş
Department of Electronics and Communication Engineering
Yıldız Technical University
İstanbul, Türkiye
eren.okumus@std.yildiz.edu.tr

4th Ömer Faruk Gülhan
Department of Electronics and Communication Engineering
Yıldız Technical University
İstanbul, Türkiye
faruk.gulhan@std.yildiz.edu.tr

*5th Özden Niyaz
Department of *Electronics and Communication Engineering*
Yıldız Technical University
İstanbul, Türkiye
oniyaz@yildiz.edu.tr

Keywords— waste management, artificial intelligence based systems, deep learning, image recognition, route optimization

I. EXTENDED ABSTRACT

This study presents a novel hybrid waste management system that integrates sensor-based data collection, vision-based inspection, image processing, and artificial intelligence (AI)-based classification. To address the limitations of traditional waste collection methods, which typically rely on fixed schedules, a prototype system was designed and simulated on the Yıldız Technical University Davutpaşa Campus. The prototype aims to determine the fill levels of waste containers as accurately as possible, reduce unnecessary collection trips, and optimize vehicle routing. By doing so, the system helps decrease operational costs, fuel consumption, and environmental impact, thus contributing to more sustainable campus-wide waste management [1-3]. The core innovation lies in combining ultrasonic sensor measurements with vision-based validation and classification models to detect fill levels and ensure that waste containers are collected at the right time. System overview is shown in Fig.1.

Fig. 1. Hybrid waste management system overview.

For sensor-based data collection, the HC-SR04 ultrasonic sensor was chosen to provide accurate distance measurements while mitigating common issues with infrared sensors, such as susceptibility to ambient light or heat. The system is designed to place each HC-SR04 sensor at a specific position on the waste container, allowing continuous detection of the trash level. The sensor data are transmitted wirelessly via the nRF24L01 module to a Raspberry Pi-based processing unit. The wireless communication component ensures that data from multiple containers can be gathered effectively without needing extensive wiring, which can be both costly and cumbersome for large-scale deployments. The Raspberry Pi processes incoming sensor data collected from trash bins with varying fullness capacities. Subsequently, a simple regression model is applied on the computer to estimate the percentage fill level of each container. The output of this regression process is then stored in a cloud database for real-time access and further decision-making [4]. This step makes it possible to integrate results from multiple sensors across the campus and update the waste management system's central control dashboard instantly.

An AI-driven image classification mechanism complements the sensor data by detecting anomalies or errors in the ultrasonic readings, enabling visual confirmation using image processing techniques [5-7]. In this hybrid approach, each container can be photographed to verify sensor output when necessary, especially if the sensor data exhibit unusual patterns or sudden changes. The image processing pipeline involves basic operations such as resizing, noise reduction, and contrast enhancement, ensuring that the inputs are suitably clear for classification. In future implementations, drone-based imaging can further support this validation process by flying over multiple container locations and capturing images without requiring manual intervention. This drone-assisted approach can be especially advantageous in large or hard-to-reach areas, improving both efficiency and safety for the personnel involved in waste management [8,9].

To enhance waste classification, a deep learning approach leveraging the YOLO object detection algorithm was employed [10]. A custom dataset of 150 labeled images was compiled, consisting of 75 images labeled as "Empty" and 75 labeled as "Full," captured from various angles and fill levels as shown in Fig. 2. These images reflect typical conditions found on the university campus, including variations in lighting and slight obstructions such as partially covered containers. The collection of a custom dataset was necessary because publicly available waste container image datasets did not fully match the specific conditions or label structure required for this project. Once assembled, the images were used to train a YOLOv8 model, which was selected due to its ability to provide rapid detection and classification in real-time scenarios. The model's output which has shown in Fig. 3. is integrated into the system's decision-making pipeline, effectively cross-verifying the sensor-reported fill levels with visual evidence to minimize false positives or negatives.

Fig. 2. Samples from the collected image dataset, showcasing labeled waste container images used for AI-based classification and fill-level detection.

This classification output then informs the route optimization module, which applies Dijkstra's shortest path algorithm to determine the most efficient collection paths [11]. By incorporating both sensor data and image validation, the routing system adapts to real-time conditions and identifies which containers are genuinely at or near full capacity. As a result, collection vehicles only visit those containers that require immediate attention, significantly reducing unnecessary trips. This optimization not only saves fuel but also reduces vehicle wear and tear, contributing to lower maintenance costs over time. The reduction in vehicle movement further aligns with broader environmental goals, such as minimizing carbon emissions and improving air quality on campus.

Experimental results presents that the proposed hybrid approach will significantly reduce trip duration and fuel usage. In comparison with traditional fixed-schedule collection routines, the system's real-time tracking of container fill levels resulted in fewer overall trips. Operational data indicate a noticeable drop in both total distance traveled and average fuel consumption per trip, validating the economic and environmental benefits of the approach. Additionally, the integration of sensor-based measurements with vision-based classification proved highly effective in detecting anomalies. For example, if the ultrasonic sensor incorrectly reports an unusually high fill level due to temporary obstructions, the imaging component can correct this reading, preventing unnecessary dispatch of a collection vehicle.

This study contributes to the broader smart city agenda by combining multiple technologies—sensor data, image processing, and AI-based classification—to improve reliability and efficiency in urban waste management. The use of cloud-based data storage further facilitates real-time monitoring and data-driven decision-making. Future work includes expanding the dataset, refining the AI model for diverse environmental conditions such as nighttime or harsh weather, and scaling the solution to larger urban areas. With a larger and more varied dataset, the detection and classification model can become more robust, improving accuracy in different scenarios. The system architecture can also be adapted to incorporate additional features like predictive maintenance for waste containers, automated alerts to maintenance crews, or integration with existing campus security drones. These findings align with existing research in smart waste management and highlight the potential of AI-driven solutions in advancing urban sustainability. By leveraging this hybrid approach, municipalities and campus administrators can ensure a more responsive, cost-effective, and eco-friendly waste management process, ultimately laying the groundwork for future innovations in the field.

Fig. 3. Image recognition result.

Acknowledgment

This work was supported by The Scientific and Technological Research Council of Türkiye (TÜBİTAK) (Project Type:2209-B University Students' Industry-Oriented Research Projects Support Program, Project Submission Number: 1139B412202746, Project Term: 2-2022.)

References

[1] N. Min-A and S. Alrashed, "Smart Campus—A Sketch," *Sustainable Cities and Society*, vol. 59, 2020. Available: https://www.sciencedirect.com/science/article/pii/S2210670720302183. [Accessed: Nov. 9, 2022].

[2] P. A. Lokhande and M. D. Pawar, "A Review: Garbage Collection Management System," *International Journal of Research in Engineering and Technology*. Available: https://www.academia.edu/28689051/A_REVIEW_GARBAGE_COLLECTION_MANAGEMENT_SYSTEM. [Accessed: Jan. 6, 2024].

[3] S. A. Ghadage and N. A. Doshi, "IOT-Based Garbage Management (Monitor and Acknowledgment) System: A Review," in *Proc. 2017 Int. Conf. Intell. Sustainable Syst. (ICISS)*, 2017. doi: 10.1109/iss1.2017.8389250. [Accessed: Jan. 6, 2024].

[4] T. Catalina, J. Virgone, and E. Blanco, "Development and Validation of Regression Models to Predict Monthly Heating Demand for Residential Buildings," *Energy and Buildings*, vol. 40, no. 10, pp. 1825–1832, 2008. doi: 10.1016/j.enbuild.2008.04.001. [Accessed: Jan. 6, 2024].

[5] T. Gupta et al., "A Deep Learning Approach-Based Hardware Solution to Categorize Garbage in the Environment," *Complex & Intelligent Systems*, vol. 8, no. 2, pp. 1129–1152, 2021. doi: 10.1007/s40747-021-00529-0. [Accessed: Jan. 6, 2024].

[6] A. Altikat, A. Gulbe, and S. Altikat, "Intelligent Solid Waste Classification Using Deep Convolutional Neural Networks," *International Journal of Environmental Science and Technology*, vol. 19, no. 3, pp. 1285–1292, 2021. doi: 10.1007/s13762-021-03179-4. [Accessed: Jan. 6, 2024].

[7] S. Majchrowska et al., "Deep Learning-Based Waste Detection in Natural and Urban Environments," *Waste Management*, vol. 138, pp. 274–284, 2022. doi: 10.1016/j.wasman.2021.12.001. [Accessed: Jan. 6, 2024].

[8] T. Filkin, N. Sliusar, M. Ritzkowski, and M. Huber-Humer, "Unmanned Aerial Vehicles for Operational Monitoring of Landfills," *Drones*, vol. 5, no. 4, p. 125, 2021. doi: 10.3390/drones5040125. [Accessed: Jan. 6, 2024].

[9] M. Z. Mat Azmi, M. A. Mohamad Sharom, S. M. Md Zin, W. Numpang, and A. R. Sipit, "Landfill Identification through High-Resolution Satellite Images and Verification Using Drone Full Motion Video (FMV)," *ASM Science Journal*, vol. 17, pp. 1–11, 2022. doi: 10.32802/asmscj.2022.823. [Accessed: Jan. 6, 2024].

[10] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.

[11] M. Noto and H. Sato, "A method for the shortest path search by extended Dijkstra algorithm," in *Proc. 2000 IEEE Int. Conf. Systems and Cybernetics*, vol. 3, 2000. doi: 10.1109/ICSMC.2000.886462.

SIMULATION FRAMEWORK FOR REAL-TIME MANAGEMENT AND CONTROL OF MANUFACTURING SYSTEMS

Iwona ŁAPUŃKA[1], Piotr WITTBRODT[2]

[1]Opole University of Technology, Sosnkowskiego 31, 45-272 Opole, Poland, i.lapunka@po.edu.pl

[2]Opole University of Technology, Sosnkowskiego 31, 45-272 Opole, Poland, p.wittbrodt@po.edu.pl

Abstract
The aim of this paper is to develop the assumptions of a simulation model for discrete events realised in heterogeneous manufacturing systems under deterministic constraints. The representation of the manufacturing system under study in the form of a simulation model will allow the determination of the system's conditions for a given workstation and the average number of items in the queue, and will enable its limitations to be analysed under real conditions. A generic simulation model was constructed based on a mathematical model of the resource-constrained project scheduling problem (RCPSP) with defined key scheduling parameters and evaluation metrics. The case study presented in this paper incorporates all the adjustments required to apply the proposed model. An example of the application of simulation-based optimization will be demonstrated in the FlexSim environment. This paper is the result of the authors' broader research work related to decision-making systems using distributed intelligence and the resulting necessary automatic organisational changes in discrete manufacturing systems.

Key Words: *manufacturing system; RCPSP model; soft computing; computer simulation; FlexSim simulation environment*

1. Introduction

The challenges currently faced by most manufacturing companies are the result of the diversification strategy being adopted in the market, which leads to an increase in the range on offer, while at the same time diversifying order volumes. Currently, due to specific customer requirements, the majority of manufacturing is carried out as make-to-order (MTO) and only part of the product range, including mainly standard, fast-moving products, is produced as make-to-stock (MTS). Flexible complexes of operations conditioned logically in time and space, with a variable structure adapted to the quantitative and qualitative characteristics of the manufactured products, correspond to discrete processes dominated by discontinuous energy and materials flows. Discrete manufacturing is now the predominant type of production in Poland. Dedicated to the furniture, automotive and electronics industries, among others, it requires the preparation of complex engineering documentation as well as control programmes for manufacturing equipment and complex production planning. The biggest problem in managing multi-range production with mostly very limited resources is the high variability of both internal and external conditions. Generating optimal operational schedules in heterogeneous systems under deterministic constraints is impossible, resulting in sub-optimal solutions that, taking into account conflicting optimisation criteria, satisfy to the expected extent the market (maximisation of customer service levels) and production (minimisation of costs) objectives of modern industrial companies is becoming increasingly desirable.

The increasing complexity of problems in production planning and control, coupled with the difficulties in finding solutions through mathematical models within a reasonable timeframe, has led to a growing interest in soft computing methods [1-3]. This is particularly relevant to integrating artificial intelligence and machine learning methods with simulation techniques, especially in developing cyber-physical systems powered by simulation models [4]. This direction of development holds significant potential, and manufacturers and designers of modern production systems are gaining increasing interest in such solutions. The simulation results and visualisations can be useful, among others, in analysing the impact of the timetable and capacity of transport vehicles, the frequency of material supply, the location and capacity of warehouses and the logic of handling production resources on the ability to meet manufacturing norms. They also allow for the detection of constraints in current conditions, possible collision risks and experiments that take into account the possibilities of synchronising different flows at the plant.

2. Literature review

The potential for the application of simulation and visualization in production planning and control is increasing alongside the development of advanced simulation software and the growing capability to easily create models with a high degree of fidelity in representing production systems at various levels and their operational dynamics. In practice, simulation-assisted planning and control systems often necessitate that responses generated by the simulation system be available in real-time. This, in turn, imposes the requirement to develop simulation models capable of communicating with physical objects and other information systems used in manufacturing enterprises and utilizing data retrieved from these sources during the simulation process, also in real-time.

In production planning, particularly for small batch and unit production of complex products, mathematical models of the resource-constrained project scheduling problem (RCPSP) have found widespread application [5, 6]. RCPSP is a deterministic problem which provides an extension to the critical path method (CPM) adding resource usage constraints [7]. A project should be viewed as the execution of multiple production orders, each of which could be described as an independent project. However, project-orders executed by means of apparent (fictitious) operations and sequence constraints can be combined into a single project (portfolio), describing the implementation of the entire production plan [8]. For an overview of the models and algorithms used in the RCPSP problem, see, for example, the works of Demeulemeester and Herroelen [9], as well as Hartmann and Briskorn [10].

When developing a production plan, the aim should be to ensure that the resource demand equals their availability, to limit their fluctuation and to eliminate over-utilisation within a given planning horizon. The delivery time for the entire project (as well as individual processes) depends on resource availability. Typically, an increase in the number of resource units available is associated with an increase in the cost of project implementation. Equalisation of employment can also be achieved through the time-varying performance intensity of non-critical processes.

RCPSP model is extensively discussed in the literature and has well-documented applications in the machinery industry, shipbuilding, and construction sectors. In such manufacturing processes, the number of customer orders (component projects) is small, so combining them into batches is either easy or impossible, the products consist of many parts and sub-assemblies, and the sequencing constraints of the tasks are very complex, the number of operations can run into the hundreds or even thousands. In these companies, many workers are qualified for various jobs, and the machines are not overloaded, so one can then be satisfied with an approximate and aggregate consideration of limited resources.

RCPSP models assume that the number and batch sizes of all products and semi-finished products are given, thus the number and performance times of all operations are known. The complex structure of products in this case equates to complex sequencing constraints between operations. The timing of operations in a RCPSP task should be set in such a way as to minimise the selected criteria, usually the entire project's completion time. Unlike classic project planning tasks, in RCPSP tasks different resources, workers, machines or tools may be needed to perform each operation. The timing of operations should be planned so that the total load on a resource never exceeds its availability [11].

RCPSP/max models allow for mathematical descriptions of a variety of deadline constraints that are generally only described verbally in scheduling theory (ST) models, e.g. the performance of several consecutive operations on the same product or on the same machine without interruption, or the simultaneous start or end of several operations, which can be useful for describing production processes in the chemical and pharmaceutical industries and are therefore also used for short-term scheduling. RCPSP models and algorithms can be considered as part of the scheduling theory, but project scheduling methods have always been seen as a separate field.

3. Problem statement and methods

A generic simulation model was constructed based on a mathematical model of the resource-constrained project scheduling problem (RCPSP) with defined key scheduling parameters and evaluation metrics that take into account both the performance measures used for deterministic problems (including cost-effectiveness) and the resilience and stability measures specific to task scheduling under uncertainty. The model's implementation in the FlexSim simulation environment allowed its effectiveness to be tested for heterogeneous manufacturing systems executing make-to-stock (MTS), make-to-order (MTO) and engineer-to-order (ETO) tasks concurrently. Ultimately, the overall simulation model was refined by successively selecting the experiment's key parameters adequately to the varying conditions of manufacturing execution in a variety of production environments.

The RCPSP seeks minimize the project duration taking into account the availability of resources, the activities duration and their precedence. In the formulation of the optimization model *n* activities are defined and identified by a subscript (*i* or *j*) with values between 1 and *n*. For practical programming reasons, the first and last activity are represented as fictitious activities of zero duration.

The time horizon is divided into partial periods represented by a subscript (*t*) with values between 1 and *T*. In each period, activity *i* consumes r_{ik} units of the renewable resource *k*. The maximum amount of renewable resource *k* available for each period *t* is defined as W_{kt}. A mathematical model to solve the RCPSP can be expressed as linear program [12]:

$$Minimize\ S_n \tag{1}$$

Subject to:

$$S_j \geq S_i + d_i \qquad \forall\ (i,j) \in E \tag{2}$$

$$\sum_{i \in A(t)} r_{ik} \leq W_{kt} \qquad \forall\ k, \forall\ t \tag{3}$$

$$S_i \geq 0 \qquad \forall\ i \tag{4}$$

Where S_i represents the start times of each activity *i*, and d_i its duration. Eq. (2) establishes the start times for each activity *j* taking into account the start times and the duration of the preceding activities. *E* represents the set of activities and their precedences. Eq. (3) refers to the restriction of renewable resources, since the active activities in period *t* cannot exceed the available amount of type *k* resources. Some mathematical models also include non-renewable resources and partially renewable resources [13]. Finally, eq. (4) indicates that the start times are continuous positive variables. Under a non-deterministic focus, the input parameters of the integer linear program (d_i, r_{ik}, or W_{kt}) must be modeled as probability functions, fuzzy numbers or other type of representation [14].

4. Results and discussion

Due to the different way in which the organisational structure of project-type orders is modelled in manufacturing companies, they are classified as two basic types: operations complex-type projects and projects that can be organised by direct line methods of work organisation. Operations-complex type projects involve technologically heterogeneous processes that are not characterised by cyclical and rhythmic performance.

The algorithms developed so far focus on finding optimal solutions in RCPSP models with the common objective of reducing waiting times, i.e. minimising losses in production. A broader view of the problem, based on process and project management knowledge, reveals the need to assess the profitability of a process/project in conjunction with the reduction of waiting times: (1) for the customer, (2) for the product, (3) for the machine. The hardest issue to decide is which waiting time to reduce in order to generate the greatest savings. The level of difficulty arises from the dynamics of the process itself and the impact of these three waiting times on the total cost of the process. This type of approach in the area of decision-making is defined by the trade-off relationship, which belongs to the set of NP-hard problems. This is one of the hitherto unexplored areas of operations research and is a common decision-making problem in today's dynamic manufacturing systems. An additional innovative approach is the use of discrete-event simulation as a tool for analysing processes with a stochastic flow structure, in which the relationships between the competing costs associated with customer, machine and product waiting times play an important role.

The case study presented in this paper incorporates all the adjustments required to apply the proposed model. An example of a project involving two orders is shown in Fig. 1. The first order (green line) requires operations 1, 2, 3, 4, 5 and 6, and the second order (red line) requires operations 7, 8, 9, 10 and 11. Operation 0 (start) is linked by apparent relations to the start operations of both orders, and the end operations of both orders are linked to the apparent end relation of the whole plan 12 (end). The operations' duration and resource requirements have been recorded as simulation data in FlexSim, assuming that critical project resources may restrict the simultaneous scheduling of certain activities.

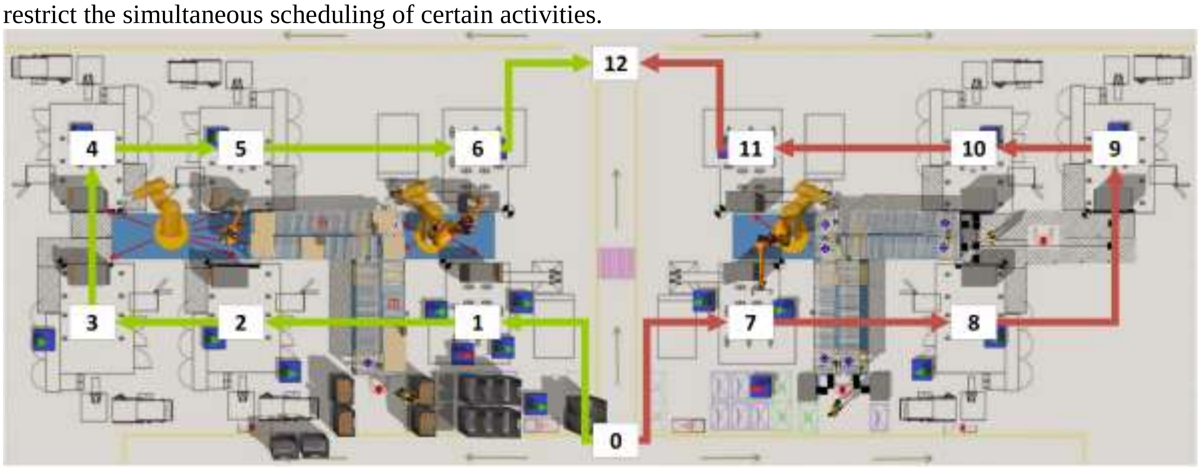


Fig. 1. Example network diagram for RCPSP in the production system

Procedures to solve the RCPSP with a random duration of activities are designed considering predictive, reactive, or proactive strategies. According to [15], a predictive strategy takes the average durations of the activities as input data to create a project baseline. A reactive strategy re-schedules the original baseline when an unexpected event occurs. A proactive strategy considers the variation of the activity duration to generate a robust baseline for the project. The baseline can be regarded as strong when few adjustments are required to face the project's disruptions. Based on the tested scenarios, various simulation results were obtained. The selection of the optimal scenario is determined by the evaluation criteria adopted for a given case. A sample simulation result for operations 1 and 2 is presented in Fig. 2.

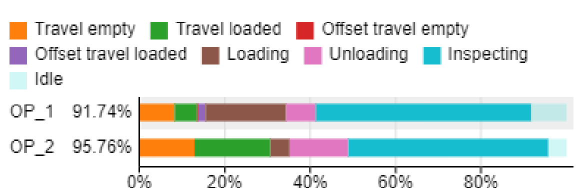


Fig. 2. Simulation results for operations 1 and 2

Resource-constrained project scheduling problem (RCPSP) models assume that the number and size of batches are known, but struggle to describe the various resource constraints or machine configurations. However, RCPSP algorithms can easily cope even with numerous operations and very complex sequence constraints.

Using the RCPSP model, which allows for defining maximum intervals between operations (RCPSP/max) rather than just minimum intervals, it is possible to describe a number of specific conditions occurring in the manufacturing process [16], e.g. the division of a production batch into transport batches, the requirement to start or finish a group of operations simultaneously, the need to perform successive operations of a given product or machine without interruption. This relatively simple model allows therefore for describing many complex production issues, but not all. New special application cases are still being considered, including: random performance times, various types of changeovers, variable operation length [17], common planning [18], multi-skilled workers [19], activities that require discrete and continuous resources simultaneously [20]. Unfortunately, solving RCPSP tasks is not a simple matter, so in addition to the established methods [5, 6], new methods, metaheuristics [17, 21] and heuristics MIP (mixed integer programming) [19] continue to appear in the literature.

5. Conclusions

The application of simulation techniques in production planning and control processes is closely linked to the concept of dynamic models (data-driven dynamic simulation) and symbiotic simulation systems. The notion of symbiotic simulation, which has been increasingly discussed in academic literature over the past two decades, is defined as a tight integration between simulation and production systems. The most frequently mentioned example of symbiotic simulation system applications is the so-called digital twin of a production system or subsystem based on simulation models [3].

Solutions based on simulation models, combined with the capabilities of modern data acquisition and transmission systems, as well as, more broadly, contemporary IT technologies, can serve as effective tools to support production management at the operational level. Each production system component can be modeled with data assigned to it and updated in real-time based on data from the actual objects within the system, their operational logic, and interconnections with other objects.

Designing a production planning system for any company is a major challenge. No known model is perfect, as none of them describe all the relevant process features and known algorithms perform better or worse under different conditions. Very detailed modelling of the production process and very detailed planning are not always necessary in manufacturing practice. Many times, decision-makers choose acceptable solutions (including Pareto-optimal solutions) because of the rate at which they are generated, while satisfying the boundary and limiting conditions for a particular objective function. However, in each case, making a choice requires knowledge of both the model features and the algorithm effectiveness, as well as an in-depth knowledge of the manufacturing process itself, together with the implementation environment.

Acknowledgements

This work was financially supported by the Opole University of Technology as part of the DELTA project no. 268/24.

References

[1] Gunal M, editor. *Simulation for Industry 4.0: past, present, and future*. 1st ed. Cham: Springer Nature Switzerland AG; 2019.
[2] Mourtzis D, Simulation in the design and operation of manufacturing systems: state of the art and new trends. *Int J Prod Res* 2020, **58**(7):1927–1949.
[3] Tliba K, Diallo TML, Penas O, Khalifa RB, Yahia NB, Choley JY, Digital twin-driven dynamic scheduling of a hybrid flow shop. *J Intell Manuf* 2023, **34**:2281–2306.
[4] Krenczyk D, Pawlewski P, Plinta D. *Symulacja procesów produkcyjnych*. Warszawa: Polskie Wydawnictwo Ekonomiczne; 2022.
[5] Kolisch R. *Project scheduling under resource constraints. Efficient heuristics for several problem classes*. Berlin-Heidelberg: Springer-Verlag; 1995.
[6] Neumann K, Schwindt C, Zimmermann J. *Project scheduling with time windows and scarce resources*. 2nd ed. Berlin-Heidelberg: Springer-Verlag; 2003.
[7] Artigues Ch, Demassey S, Néron E, *Resource-constrained project scheduling: models, algorithms, extensions and applications*. London-Hoboken: ISTE-Wiley; 2008.
[8] Hartmann S, *Project scheduling under limited resources: models, methods, and applications*. 1st ed. Berlin-Heidelberg: Springer-Verlag; 1999.
[9] Demeulemeester EL, Herroelen WS. *Project scheduling. A research handbook*. Dordrecht: Kluwer Academic Publishers; 2002.
[10] Hartmann S, Briskorn D. A survey of variants and extensions of the resource-constrained project scheduling problem. *Eur J Oper Res* 2012; **207**(1):1–14.
[11] Kaczmarczyk W. Przegląd zaawansowanych modeli planowania produkcji. In: *Zarządzanie przedsiębiorstwem. Teoria i praktyka*, Kraków: Wydawnictwa AGH; 2015, p. 39–52.
[12] Palpant M, Christian A, Michelon P. LSSPER: solving the resource-constrained project scheduling problem with large neighbourhood search. *Ann Oper Res* 2004, **131**:237–257.
[13] Böttcher J, Drexl A, Kolisch R, Salewski F, Project scheduling under partially renewable resource constraints. *Manage Sci* 2008, **45**(4):543–559.
[14] Ortiz-Pimiento NR, Diaz-Serna FJ. An optimization model to solve the resource constrained project scheduling problem RCPSP in new product development projects. *DYNA* 2020; **87**(212):179–188.
[15] Brčić M, Kalpić D, Fertalj K, Resource constrained project scheduling under uncertainty: a survey. In: *23rd Central European Conference on Information and Intelligent Systems*; 2012, p. 401–409.
[16] Neumann K, Schwindt C. Activity-on-node networks with minimal and maximal time lags and their application to make-to-order production. *OR Spektrum* 1997; **19**:205–217.
[17] Józefowska J, Węglarz J, editors. *Perspectives in modern project scheduling*. New York: Springer; 2006.
[18] Pinto G, Ben-Dov YT, Rabinowitz G. Formulating and solving a multi-mode resource-collaboration and constrained scheduling problem (MRCCSP). *Ann Oper Res* 2013; **206**(1):311–339.
[19] Montoya C, Bellenguez-Morineau O, Pinson E, Rivreau D. Branch-and-price approach for the multi-skill project scheduling problem. *Optim Lett* 2014; **8**(5):1721–1734.
[20] Waligóra G. Discrete-continuous project scheduling with discounted cash inflows and various payment models - a review of recent results. *Ann Oper Res* 2014; **213**(1):319–340.
[21] Diakoulakis IE, Koulouriotis DE, Emiris DM. Resource constrained project scheduling using evolution strategies. *Oper Res* 2004; **4**(3):261–275.

Lung and Colon Cancer Histopathological Image Classification Using Deep Learning Approaches

Oussama Bouguerra[1], Youcef Brik[2], Bilal Attallah[3]

[1]*LASS Laboratory Department of Electronics University of M'sila* **M'sila, Algeria, oussama.bouguerra@univ-msila.dz**
[2]*LASS Laboratory Department of Electronics University of M'sila* **M'sila, Algeria youcef.brik@univ-msila.dz**
[3]*LASS Laboratory Department of Electronics University of M'sila* **M'sila, Algeria, bilal.attallah@univ-msila.dz**

Abstract—**This paper investigates the application of deep learning for multi-class classification of lung and colon cancer histopathological images using the LC25000 dataset. The dataset contains 25,000 images equally distributed across five classes: benign and adenocarcinoma for both lung and colon tissues, and squamous cell carcinoma for lung tissue. Preprocessing involved converting labels to one-hot encoding and removing 1280 identified duplicate images to prevent data leakage. Three models were trained and evaluated: a baseline CNN, an enhanced CNN with architectural modifications and regularization, and a transfer learning model utilizing ResNet50. The baseline model, using Adam optimizer and categorical cross-entropy loss, achieved a test accuracy of 62.6% and validation accuracy of 63.7%. The enhanced model, incorporating increased depth, adjusted kernel size, L2 regularization, dropout, and the Adamax optimizer with early stopping, reached a test accuracy of 97.9% and validation accuracy of 98.1%. Finally, the transfer learning model with ResNet50, fine-tuned with additional dense layers, dropout, and early stopping, achieved near-perfect performance with 98.9% test accuracy and 99.3% validation accuracy. This study demonstrates the significant performance improvement achieved through architectural enhancements and transfer learn- ing, highlighting the potential of deep learning for automated diagnosis of cancer from histopathological images. The removal of duplicate images proved essential for accurate performance evaluation and preventing artificially inflated results due to data leakage.**

***Index Terms*—Histopathological Image Classification,Lung and Colon Cancer,Convolutional Neural Network (CNN),Transfer Learning,ResNet50.**

I. Introduction

Lung and colon cancers are among the leading causes of cancer-related mortality worldwide, underscoring the crit- ical need for early detection to improve patient outcomes. Histopathological imaging plays a pivotal role in the diagnosis of these cancers, as it allows for the detailed examination of tissue samples to identify malignant cells. However, the high variability in tissue structures and staining techniques presents significant challenges in accurately interpreting these images. Studies have highlighted the importance of histopatho- logical examination in colorectal cancer, emphasizing that early detection can significantly expand treatment options and reduce mortality rates [1]. Similarly, the manual analysis of histopathology images for lung cancer detection is labor- intensive and highly dependent on the pathologist's expertise, which can lead to variability in diagnostic accuracy [2]. These challenges necessitate the development of automated, reliable methods to enhance the precision and efficiency of cancer diagnosis. Deep learning, particularly convolutional neural net- works (CNNs), offers promising solutions to these challenges by improving the accuracy of histopathological image analysis. CNNs have demonstrated impressive capabilities in image pro- cessing tasks, including medical imaging, by learning complex patterns and features from large datasets [2]. Transfer learning, using architectures like ResNet, has been particularly effective in leveraging pre-trained models to enhance classification performance, even with limited labeled data [1], [2]. This approach allows for the adaptation of models trained on large, diverse datasets to specific tasks, such as cancer detection, thereby improving generalization and reducing the need for extensive annotated datasets [3]. Despite these advancements, challenges remain, such as the need for models to generalize across different centers and staining protocols, which can lead to overfitting if not properly addressed [4]. Nonetheless, the integration of deep learning techniques in histopathological image analysis holds significant potential for advancing cancer diagnostics and improving patient care.

II. Related Work

Convolutional Neural Networks (CNNs) have revolution- ized medical imaging, offering significant advancements in classification, segmentation, and detection tasks. Seminal ar- chitectures like AlexNet, VGG, and ResNet have laid the groundwork for these applications, demonstrating their effi- cacy in various medical domains. For instance, CNNs have been pivotal in enhancing diagnostic accuracy and efficiency in radiology and pathology, where they have been used to automate the analysis of complex medical images, thus reduc- ing the reliance on manual interpretation and improving early disease detection

[5]. In the context of cancer histopathology, CNNs have been employed to classify lung and colon cancer images, leveraging their ability to learn intricate patterns from large datasets.

Recent studies have focused on applying deep learning, particularly CNNs, to classify cancer from histopathological images, with a strong emphasis on lung and colon cancer

For example, the BIC-SGODL method utilizes DenseNet and convolutional LSTM to capture complex features and spa- tiotemporal information, achieving high performance in cancer diagnosis [6]. Another study compared various pre-trained models, including VGG-16 and ResNet, on the LC25000 dataset, achieving accuracies up to 100% for certain cate- gories, highlighting the effectiveness of transfer learning in this domain [2]. Additionally, novel architectures like the 1D Convolutional Channel-based Attention Networks have been proposed, achieving state-of-the-art performance with minimal computational resources [7]. These studies underscore the potential of deep learning models to significantly improve the accuracy and efficiency of cancer diagnosis.

Transfer learning has emerged as a powerful technique in histopathology, allowing models to leverage pre-trained weights from large datasets to improve performance on spe- cific tasks. For instance, the Inception-ResNetV2 model, com- bined with local binary pattern features, achieved 99.98% accuracy in lung and colon cancer detection, demonstrating the utility of integrating texture-based features with deep learning [8]. Similarly, EfficientNet variants, when applied with transfer learning, achieved notable accuracy improvements, with Effi- cientNetB2 reaching 97% accuracy on the LC25000 dataset [9]. These approaches highlight the advantages of transfer learning in overcoming challenges such as limited labeled data and enhancing model generalizability.

Future research directions include developing more robust models that can handle diverse data sources and improving the interpretability of deep learning models to facilitate their integration into clinical workflows [5]. Additionally, explor- ing novel architectures and training strategies could further enhance the performance and applicability of these models in real-world settings.

III. Data Collection and Preprocessing

This section details the dataset used, preprocessing steps performed, and exploratory data analysis conducted.

A. Dataset Description

This study leverages the LC25000 Lung and Colon Cancer Histopathological Image Dataset introduced by [10], a publicly accessible resource containing 25,000 RGB histopathological images across five balanced classes: Colon Benign Tissue, Colon Adenocarcinoma, Lung Benign Tissue, Lung Adenocar- cinoma, and Lung Squamous Cell Carcinoma, with each class comprising 5,000 images. Figure 1 showcases representative images from eachclass, visually highlighting the diversity of histopathological features within thedataset.

B. Data Preprocessing

Data cleaning and preprocessing were essential steps in preparing the histopathological images for effective model training. The following methods were applied to ensure data quality, standardize input, and optimize the dataset for Con- volutional Neural Networks (CNNs):

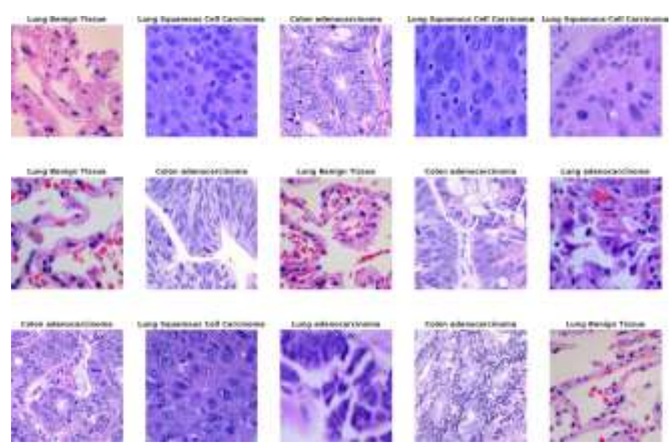


Fig. 1: Visual examples of histopathology slides from the LC25000 data

- **Data Cleaning and Duplicate Removal:** An initial exploratory data analysis (EDA) assessed class distribution and identified duplicate images. This process removed 1,280 duplicates, resulting in a final dataset of 23,720 images (See Figure 2).
- **Image Resizing**: All images were resized to 120x120 pixels using LANCZOS resampling, balancing computational efficiency with preservation of critical histopathological details for feature extraction.
- **Normalization**: Although not explicitly applied in the code, normalization is recommended to scale pixel values between 0 and 1 (or -1 and 1). This practice stabilizes model training and prevents dominant features from skewing the learning process.
- **One-Hot Encoding of Target Classes**: To support multi-class classification, categorical target variables were transformed into binary vectors using one-hot encoding, preventing unintended ordinal relationships among diag- nostic categories.
- **Train-Validation-Test Split**: The dataset was divided into training, validation, and test sets in a 60%-20%- 20% ratio, ensuring balanced class representation across all phases (Figure 3).

IV. Methodology

This section details the methodologies employed for devel- oping and evaluating the deep learning models for lung and colon cancer histopathological image classification. Figure 4 provides an overview of the proposed system, illustrating the workflow from dataset acquisition and preprocessing through model training, evaluation, and final classification.

A. Baseline Model

A baseline Convolutional Neural Network (CNN) model was established to serve as a benchmark for subsequent, more complex architectures. This model provides a founda- tion for evaluating the performance gains achieved through architectural modifications and transfer learning. The baseline model architecture, comprising convolutional, max pooling, and dense layers, is summarized in Figure 5.

B. Enhanced Model

To improve upon the baseline CNN, the enhanced model introduces additional convolutional layers, regularization tech- niques, and an optimized architecture to further capture the intricate patterns present in histopathological images. These modifications aim to reduce overfitting and enhance classifi- cation accuracy across the five cancer-related categories. The enhanced CNN architecture, incorporating increased depth, adjusted kernel size, and regularization, is detailed in Table I.

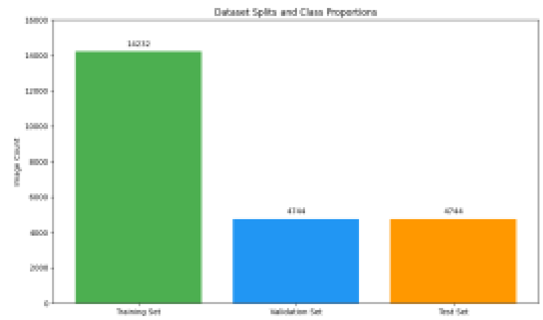


Fig. 2: Image Counts by Category Before and After Removing Duplicates.

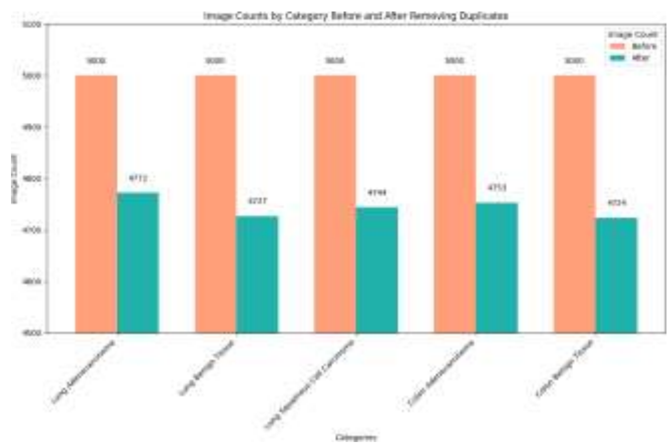


Fig. 3: Dataset Splits and Class Proportions.

C. Transfer Learning Model (ResNet50)

we employed ResNet50 a deep residual network pre-trained on the ImageNet dataset, as the foundation for our transfer learning model. The architecture of the transfer learning model, utilizing a pre-trained ResNet50 base and custom clas- sification layers, is depicted in Figure 6. This figure outlines the layer configurations and the distinction between trainable and non-trainable parameters.

D. Training Configuration

This section provides an overview of the key training configurations used across the baseline, enhanced, and transfer learning (ResNet50) models. The configurations include essen- tial hyperparameters, optimizers, learning rates, batch sizes, and epochs, highlighting the specific settings employed to achieve optimal performance for each model (see Table II).

TABLE I: Enhanced Model Architecture

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 120, 120, 64)	1792
conv2d_5 (Conv2D)	(None, 120, 120, 64)	102464
max_pooling2d_4 (MaxPooling2D)	(None, 60, 60, 64)	0
conv2d_6 (Conv2D)	(None, 60, 60, 128)	73856
conv2d_7 (Conv2D)	(None, 60, 60, 128)	147584
max_pooling2d_5 (MaxPooling2D)	(None, 30, 30, 128)	0
conv2d_8 (Conv2D)	(None, 30, 30, 256)	295168
conv2d_9 (Conv2D)	(None, 30, 30, 256)	590080
max_pooling2d_6 (MaxPooling2D)	(None, 15, 15, 256)	0
conv2d_10 (Conv2D)	(None, 15, 15, 512)	1180160
conv2d_11 (Conv2D)	(None, 15, 15, 512)	2359808
max_pooling2d_7 (MaxPooling2D)	(None, 7, 7, 512)	0
flatten_2 (Flatten)	(None, 25088)	0
dense_6 (Dense)	(None, 256)	6422784
dropout (Dropout)	(None, 256)	0
dense_7 (Dense)	(None, 64)	16448
dense_8 (Dense)	(None, 5)	325
Total params:	11190469 (42.69 MB)	
Trainable params:	11190469 (42.69 MB)	
Non-trainable params:	0 (0.00 Byte)	

V. Result

The results of this study are based on the evaluation of three deep learning models—baseline CNN, enhanced CNN, and transfer learning with ResNet50—using a diverse set of performance metrics.

A. Model Performance Comparison

Each model was assessed on accuracy, precision, recall, and F1-score, providing a comprehensive view of performance across categories.The bar graph in Figure 7 further highlights these performance improvements across models, illustrating the ResNet50 model's significant gains over the baseline and enhanced CNN models.

B. Confusion Matrix Analysis

The confusion matrices for each model illustrate the model's effectiveness in differentiating between benign and malignant tissues. Misclassifications were notably reduced in the en- hanced and ResNet50 models compared to the baseline.

C. ROC-AUC Analysis

ROC-AUC curves were generated for each model to eval- uate the models' abilities to distinguish between classes. The ResNet50 model achieved an AUC close to 1.0 across all categories, indicating strong discriminative power.

D. Discussion

The ResNet50 transfer learning model achieved the highest performance across metrics, followed by the enhanced CNN model. The baseline model, while effective as an initial benchmark, demonstrated limited accuracy and a higher rate of misclassification. The results underscore the impact of deep learning architectures and transfer learning in handling complex histopathological data, with the ResNet50 model emerging as the optimal approach for cancer classification in this study.

VI. Conclusion

This paper evaluated lung and colon cancer classification using three models: a baseline CNN, an enhanced CNN, and a ResNet50-based transfer learning model. The ResNet50 model achieved the highest performance, with 98.9% accuracy and strong precision, recall, and F1-scores, demonstrating its su- perior capability in feature extraction for complex histopatho- logical images. These results confirm the value of transfer learning in medical image analysis and set a foundation for future advancements in automated cancer diagnosis.

Acknowledgment

The authors would like to extend their gratitude to the members of LASS Laboratory (University of M'sila, Algeria) for their valuable support and assistance in publishing this work.

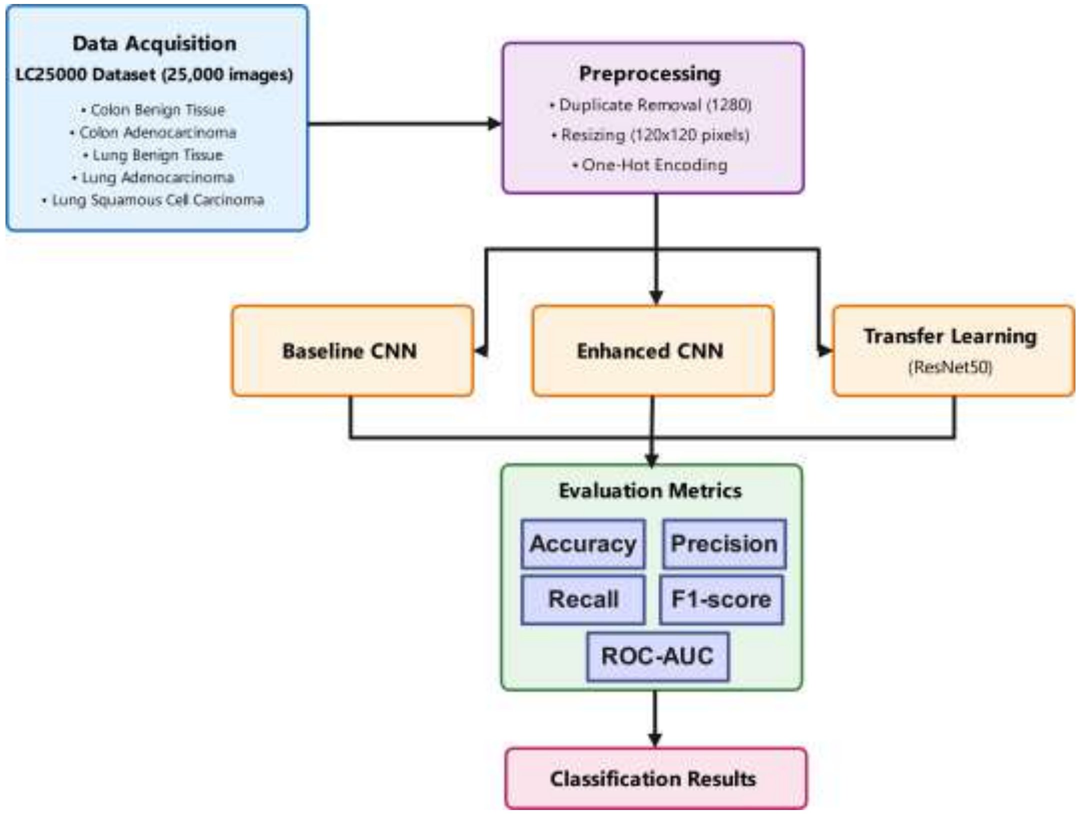


Fig. 4: Proposed System Diagram.

TABLE II: Training Configurations for Each Model

Model	Optimizer	Learning Rate	Batch Size	Epochs	Regularization	Early Stopping
Baseline Model	Adam	Default	32	10	None	No
Enhanced Model	Adamax	0.0005	32	30	L2 Regularization, Dropout (0.2)	Patience 2
Transfer Learning Model (ResNet50)	Adamax	0.0005	32	30	Dropout (0.2)	Patience 3

Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 120, 120, 128)	3584
max_pooling2d_2 (MaxPooling2D)	(None, 60, 60, 128)	0
conv2d_3 (Conv2D)	(None, 60, 60, 64)	73792
max_pooling2d_3 (MaxPooling2D)	(None, 30, 30, 64)	0
flatten_1 (Flatten)	(None, 57600)	0
dense_3 (Dense)	(None, 128)	7372928
dense_4 (Dense)	(None, 32)	4128
dense_5 (Dense)	(None, 5)	165

Total params: 7454597 (28.44 MB)
Trainable params: 7454597 (28.44 MB)
Non-trainable params: 0 (0.00 Byte)

Fig. 5: Summary of the baseline CNN architecture.

Model: "sequential_3"

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 4, 4, 2048)	23587712
global_average_pooling2d_1 (GlobalAveragePooling2D)	(None, 2048)	0
dense_9 (Dense)	(None, 256)	524544
dropout_1 (Dropout)	(None, 256)	0
dense_10 (Dense)	(None, 64)	16448
dense_11 (Dense)	(None, 5)	325

Total params: 24129029 (92.04 MB)
Trainable params: 541317 (2.06 MB)
Non-trainable params: 23587712 (89.98 MB)

Fig. 6: Architecture of the transfer learning model using ResNet50.

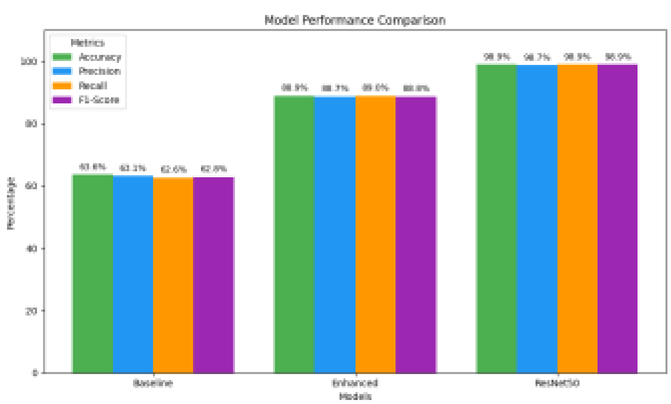
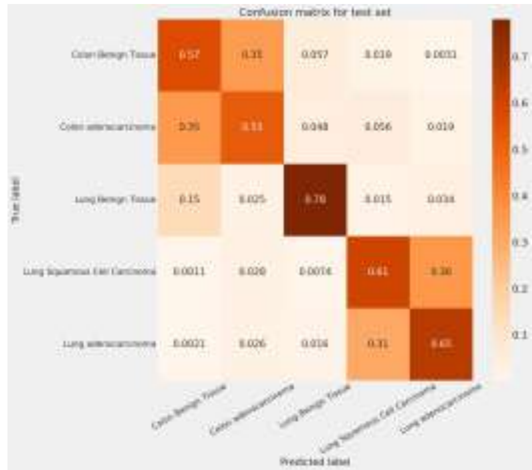
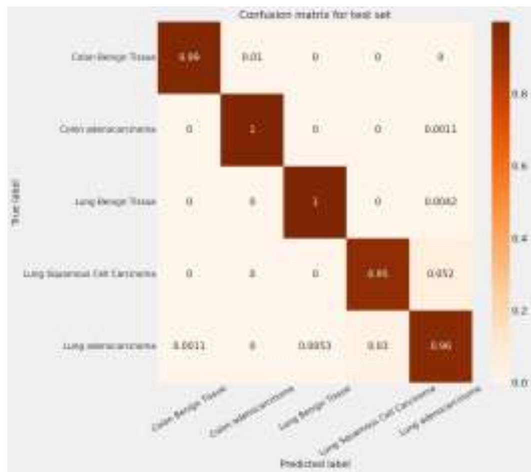


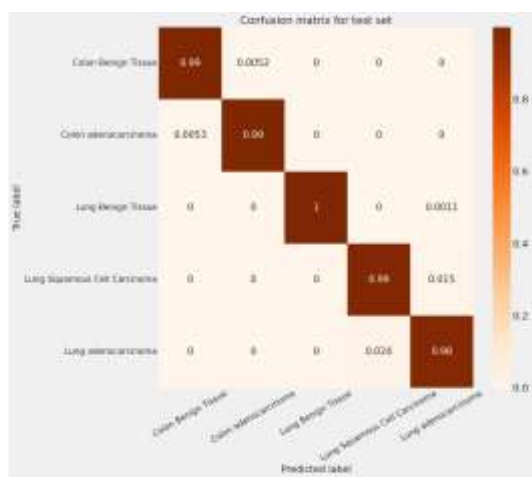
Fig. 7: Bar graph comparison of model performance metrics.



(a)

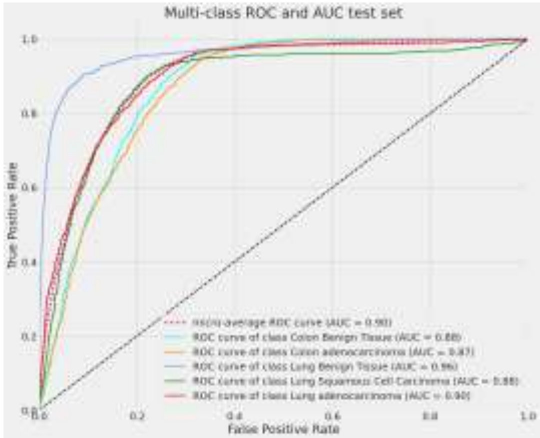


(b)

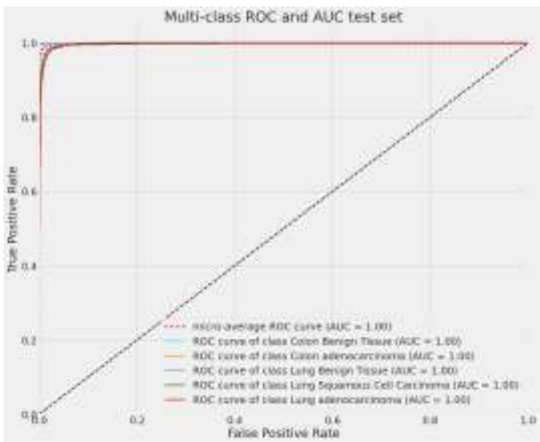


(c)

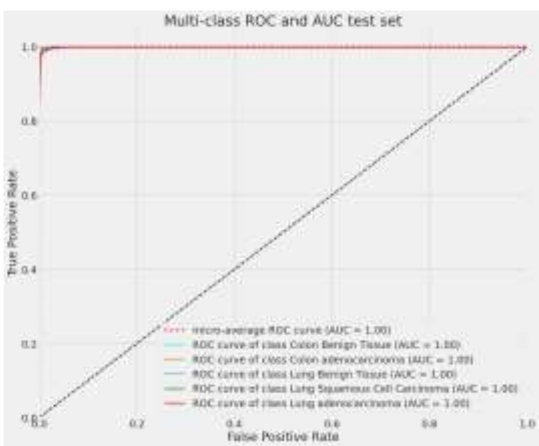
Fig. 8: Confusion matrix for (a:baseline model,b:enhanced model,c:ResNet50 model)



(a)



(b)



(c)

Fig. 9: ROC-AUC curves for (a:baseline model,b:enhanced model,c:ResNet50 model)

REFERENCES

Dina M Ibrahim, Mohammad Ali A Hammoudeh, and Tahani M Allam. Histopathological cancer detection based on deep learning and stain images. *Indonesian Journal of Electrical Engineering and Computer Science*, 36(1):214–230, 2024.

[1] Abdelwahid Oubaalla, Hicham El Moubtahij, and Nabil EL AKKAD. Detection of lung and colon cancer from histopathological images: Using convolutional networks and transfer learning. *International Journal of Computing and Digital Systems*, 16(1):583–595, 2024.

[2] Andrey Ignatov, Josephine Yates, and Valentina Boeva. Histopathological image classification with cell morphology aware deep neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6913–6925, 2024.

[3] Maryam Asadi-Aghbolaghi, Amirali Darbandsari, Allen Zhang, Alberto Contreras-Sanz, Jeffrey Boschman, Pouya Ahmadvand, Martin Ko¨bel, David Farnell, David G Huntsman, Andrew Churg, et al. Learning generalizable ai models for multi-center histopathology image classification. *NPJ Precision Oncology*, 8(1):151, 2024.

[4] Anirudh Atmakuru, Subrata Chakraborty, Oliver Faust, Massimo Salvi, Prabal Datta Barua, Filippo Molinari, UR Acharya, and Nusrat Homaira. Deep learning in radiology for lung cancer diagnostics: A systematic review of classification, segmentation, and predictive modeling techniques. *Expert Systems with Applications*, page 124665, 2024.

[5] Thiyagarajan Manoharan, Ramalingam Velvizhi, Tarun Kumar Juluru, Shoaib Kamal, Shrabani Mallick, and Ezudheen Puliyanjalil. Biomedical image classification using seagull optimization with deep learning for colon and lung cancer diagnosis. *Indonesian Journal of Electrical Engineering and Computer Science*, 2024.

[6] Nazmul Shahadat. Lung and colon cancer histopathological image classification using 1d convolutional channel-based attention networks. *The International FLAIRS Conference Proceedings*, 2024.

[7] Shtwai Alsubai. Transfer learning based approach for lung and colon cancer detection using local binary pattern features and explainable artificial intelligence (ai) techniques. *PeerJ Computer Science*, 10:e1996, 2024.

[8] Sunila Anjum, Imran Ahmed, Muhammad Asif, Hanan Aljuaid, Fahad Alturise, Yazeed Yasin Ghadi, and Rashad Elhabob. Lung cancer classification in histopathology images using multiresolution efficient nets. *Computational Intelligence and Neuroscience*, 2023(1):7282944, 2023.

[9] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wil- son, L. A. DeLand, and S. M. Mastorides. Lc25000 lung and colon histopathological image dataset, 2023. Retrieved from https://github.com/tampapath/lung$_c$olon$_i$mage$_s$et.