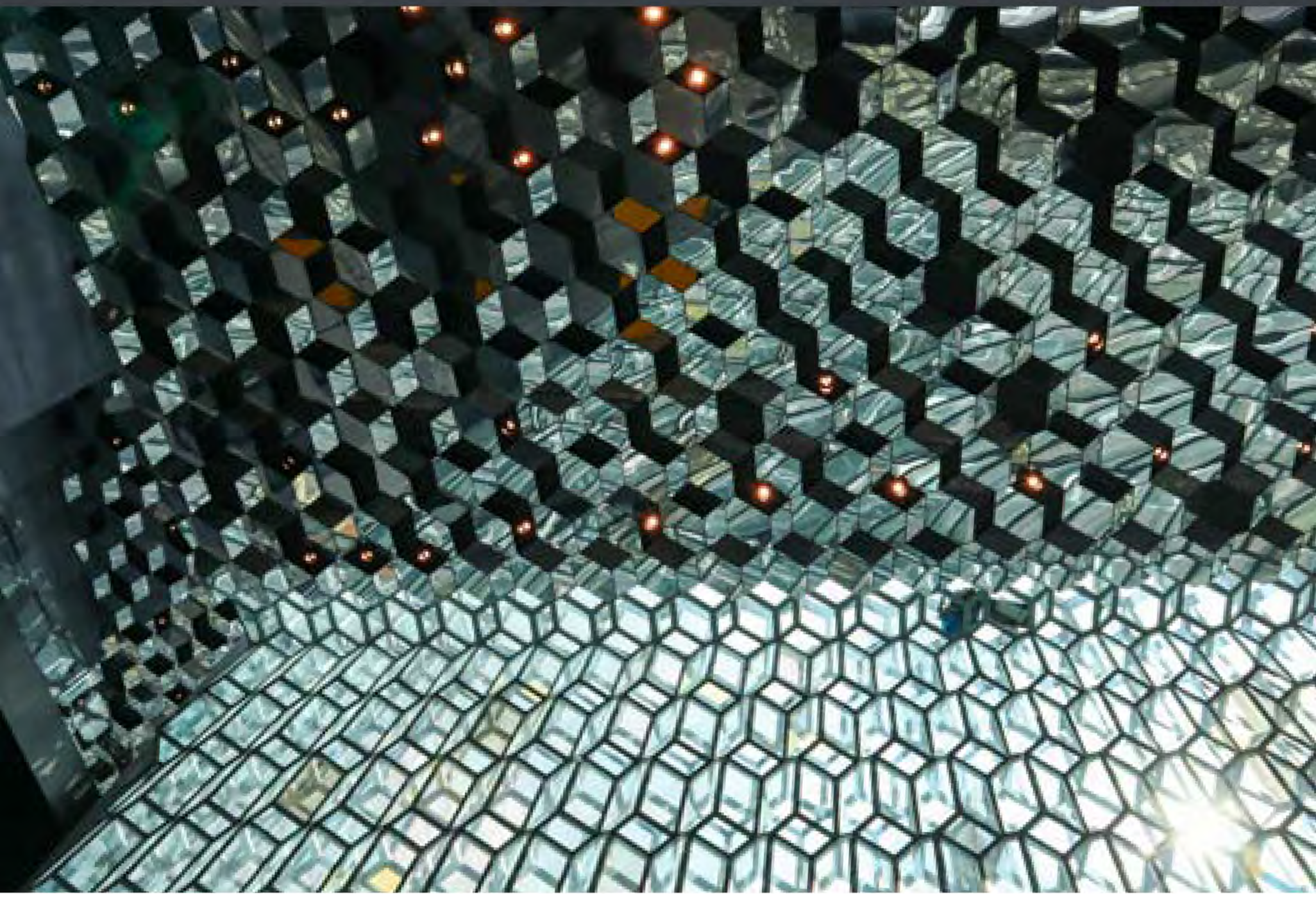
INGENIOUS

4 ISSUE 2

Editor-in-Chief: Prof. Asoc. Dr. Teuta XHINDI/ Scientific Journal of the Faculty of Engineering, Informatics and Architecture / No. 4, issue 2/ 2024

ISSN 2789-4800



INNOVATIVE TECHNOLOGIES AND SUSTAINABLE SOLUTIONS

In Engineering, Informatics, and Architecture

Prof. Asoc. Dr. Teuta **XHINDI**/ Anxhela **FEÇANJI**/ Ditmira **TAHIRI**/
Ardiana **TOPI**/ MSc. Arch. Enkeleida **PRIFTI**/ Enxhi **TAGANI**/ Erion **CURAJ**/
Flavio **KOKA**/ Joana **SHEHAJ**/ MSc. Eng. Gazmir **HALLAÇI**/ Eng. Jona **LIÇI**/
MSc. Orges **BESHKU**/ Sildi **SHAHINI**/ Forsian **ELEZI**

UETPRESS

④INGENIOUS No. 4, issue 2/2024
JOURNAL OF THE FACULTY OF ENGINEERING, INFORMATICS AND ARCHITECTURE

ISSN 2789-4800

UETPRESS

Published under the series “ACADEMIC JOURNALS”.

Disclaimer

The information and opinions presented in the Journal reflects the views of the author and not of the Journal or its Editorial Board or the Publisher.

The journal has been catalogued at the National Library of Albania and the Library of the European University of Tirana, Albania.

(print ISSN: 2789-4800/ online ISSN: 2958-888X)

ingenious@uet.edu.al
www.uet.edu.al/ingenious

UETPRESS

Published by:
EUROPEAN UNIVERSITY OF TIRANA / ALBANIA

content

EDITORIAL

Innovative Technologies and Sustainable Solutions in Engineering, Informatics, and Architecture

Prof. Asoc. Dr. Teuta XHINDI

As societies worldwide strive to balance progress with environmental stewardship, the integration of innovative technologies and sustainable solutions in engineering, informatics, and architecture has become essential. The theme "*Innovative Technologies and Sustainable Solutions in Engineering, Informatics, and Architecture*" highlights how these disciplines can lead transformative changes by adopting new technologies and fostering sustainable practices that support resilient, efficient, and responsible advancements.

In **engineering**, emerging technologies such as advanced robotics, AI, and data analytics are unlocking new pathways for sustainable solutions. For instance, studies indicate that using energy-efficient materials and methods can reduce a building's energy consumption by up to 40% (Kalbasi, Samali and Masoud, 2023), aligning with global goals to decrease carbon footprints and promote green building practices. Industrial robotics, driven by intelligent control systems, are now optimized for energy efficiency, reducing operational costs and resource use while increasing production capacity.

In **informatics**, big data and artificial intelligence offer substantial contributions to sustainable practices. The simulation findings reveal that the yearly energy consumption in the smart city may be reduced by more than 35%-40% via the optimization of energy consumption using the proposed reinforcement learning approach (Ordoueia, Broumandnia,Banirostamb and Gilanic, 2023). By using data-driven algorithms to enhance systems like transportation, water distribution, and waste management, cities can minimize waste and improve resource allocation. Research also indicates that robust cybersecurity frameworks are crucial to protect the increasing amount of data generated by these smart systems, underscoring the importance of security in sustainable digital infrastructure.

In **architecture**, the adoption of eco-friendly designs, such as ventilated facades and solar-integrated structures, has demonstrated significant environmental benefits. Several studies highlighted OVFs' positive contributions with respect to a traditional solution in summer. For example, ventilated solutions have been found to save energy in the range of 47–51 % in summer in hot Italian climates, depending on wind conditions (Gagliano, Nocera and Aneli, 2016). Architects are increasingly exploring materials that combine aesthetic value with sustainability, such as recycled composites and bio-based materials, fostering both environmental responsibility and aesthetic innovation.

This topic will serve as a platform for examining how these technologies are applied in real-world contexts. Contributors will delve into case studies, data analyses, and experimental research, such as the design and implementation of AI-driven systems for urban planning or the performance metrics of energy-efficient building materials in reducing operational costs. Through quantitative and qualitative research, this issue will highlight the tangible impact of integrating innovative technology with sustainable practices, reflecting the faculty's commitment to a forward-thinking, data-driven approach to addressing global challenges.

Also, this issue provides a rich, data-informed exploration of how the fusion of engineering, informatics, and architecture contributes to sustainable development. By presenting practical applications and empirical research, the journal aims to inspire further advancements that support a technologically advanced, eco-conscious future.

Design and Development of an Information System for the Tourist Promotion of the City of Korça

Anxhela Feçanji
European University of Tirana

Abstract

This paper explores the design and development of an innovative information system aimed at boosting tourism in the city of Korça. The problem addressed is the need for a user-centred platform that provides potential visitors the essential information while promoting the city's unique attractions. The study begins by examining relevant literature on information systems and their role in tourism, destination decision-making factors, and case studies of similar systems. It also analyzes Korça's tourism landscape and visitor preferences, identifying gaps and opportunities for innovation.

A hybrid research method combining qualitative and quantitative approaches was employed. Insights from scholarly literature informed the system's design, with a focus on engaging visitors in tourism promotion through sharing experiences and recommendations. Technologically, the system was built using HTML, CSS, and JavaScript for the front end, and MySQL, and Python (Flask Framework) for the back end, following the Model-View-Controller (MVC) architecture for ease of management and scalability.

The results demonstrate that the system effectively showcases Korça's attractions and fosters interaction between visitors and local businesses. The research concludes that this information system not only provides valuable tourism information but also enhances community engagement. Future recommendations include additional features for further improving user experience and expanding the system's reach.

Keywords: *Information System, Session, Security, Authentication, Authorization, Validation, Python, MySql, Javascript, HTML, CSS, Bcrypt, Flask, MVC Architecture, Modularization.*

1.Introduction

1.1 Field and Context of the Study

Tourism has experienced significant growth, with a 34% increase in 2023 compared to 2022, recovering to 88% of pre-pandemic levels, generating $1.286 billion in revenue (Root, 2024). It is a crucial driver of global economic development, contributing an estimated $3.3 trillion or 3% of global GDP in 2023 (UNWTO, 2024). Technological advancements, particularly in Information Technology (IT), have transformed the tourism sector, making travel safer and more accessible (Wahab, 2017). Korça, a city known for its year-round tourism potential (Zaimi, 2017), needs an effective information system to promote itself to both local and foreign visitors. Understanding consumer decision-making in tourism (Qirici, 2011) is critical for developing such a system. This study explores the role of information systems in enhancing tourism promotion and security, focusing on the design of a system tailored to Korça's visitors.

1.2 Research Question and Objectives

The primary aim of this paper is to develop an information system for promoting Korça's tourism. The research question guiding this study is:

Research Question:
How can we build an information system that provides a pleasant experience for users and includes the necessary information for potential visitors?

Objectives:

1. To understand the needs and preferences of potential visitors to Korça regarding tourist information and services.
2. To design and build an information system that meets the needs of tourists and local tourism businesses while promoting the city effectively.

1.3 Importance and Contribution of the Study

The study contributes to:

1. Tourism Promotion: By developing a user-friendly information system, it enhances Korça's tourism promotion efforts, attracting more visitors and boosting the local economy.
2. User Experience Improvement: The system focuses on delivering a pleasant experience by addressing the specific needs of potential visitors, aiming to increase satisfaction and repeat visits.
3. Academic Contribution: It enriches academic literature by examining the design and implementation of a tourism promotion information system in a specific city context.

Overall, this research is essential for promoting Korça's tourism, improving user engagement, and contributing to knowledge in the field of information systems for tourism.

1.4 Tourism in Korça

Korça, located at the foot of the Morava Mountain, is a culturally rich and important tourist destination in Albania. It is one of the country's most visited cities, attracting tourists with its medieval art museums, historical homes, churches, and natural beauty. The city also offers local cuisine and traditional drinks like "Korça" beer and vodka, with events such as the annual beer festival drawing both local and international visitors (INSTAT Albania: 2020-2022; Partners Albania, 2015).

Korça's socio-economic growth is closely tied to its natural and cultural resources, with foreign tourists drawn to its cultural heritage, including traditions and spiritual elements (Dino, 2018). Tourism in Korça is a year-round activity, with both winter and summer attractions. In the first half of 2021, the city saw over 30,000 visitors, primarily foreigners, leading to an expansion in accommodation, including modern hotels and traditional guesthouses (Xhajanka, 2021).

In a study conducted by Qirici (2011), about the reasons why tourists choose Korça as destination, it turns out that most of them (about 65.3%) were interested in natural and cultural heritage. For more, tourists mentioned obtaining information from previous visits, accommodation brochures, friends, and some of them from travel agents. The dominance of information obtained from friends or previous visits suggests that current methods of information dissemination may lack effectiveness and professionalism. The heavy reliance on informal sources, such as friends, may limit the potential for growth in the tourism market, keeping it as a casual market.

The development of a personalized information system could further enhance the tourist experience, promote local businesses, and help manage tourism sustainably.

1.5 Existing systems for promoting the city of Korça.

A thorough analysis of existing systems' functionalities is crucial before building our system. This helps identify strengths, limitations, and opportunities for innovation, allowing us to strategically enhance our system with new features that meet evolving user needs. Some of the existing systems for promoting the city of Korça are presented below.

1.Visit Korça (https://www.visit-korca.com)
This platform serves as a comprehensive information centre for Korça and surrounding areas, offering details on local attractions, monuments, cultural sites, and points of interest. It features an interactive map with historical and cultural information and provides listings of accommodations, tours, and activities in the region. Bilingual support (Albanian and English) allows access for both local and foreign tourists. However, the system has some drawbacks:

- Information is outdated (last updated in 2017).
- Incomplete map locations for hotels and attractions, making navigation difficult.

2.Discover Korça (https://discoverkorca.al/)
Discover Korça offers detailed information on hotels, restaurants, activities, art, and history, with an integrated Google Maps feature for navigation. It also provides event listings, transportation options, and the ability for users to contribute promotions. Despite being user-friendly and providing up-to-date information, it has some weaknesses:

- Issues with registration, including confusion with email-only login and failure to update new emails.
- Lack of multilingual support for tourist categories.

3.Tourist in Korça (https://turistnekorce.wordpress.com/)
This platform provides a wide range of information, from city activities to cultural centers and tourist attractions. Users can explore different categories, leave comments, and access a photo gallery for a visual tour. However, the platform has some limitations:

- Information is not always up-to-date, especially for hotels and tourist spots.
- Contact details for hotels are missing.
- Available only in Albanian, limiting its usability for foreign tourists.

2.Literature Review

2.1 The Role of Information Systems in Tourism

Over the past five decades, the travel and tourism industry has evolved into a significant component of the global economy, becoming one of the largest and fastest-growing sectors. This multifaceted industry comprises a variety of activities driven by travellers seeking unique experiences and destinations aiming to attract visitors. As a result, tourism has become a complex ecosystem, influencing cultural, social, and financial aspects of society (Sanju, 2023).

Regarding the role of tourism in economic growth, Sanju (2023), emphasizes that this sector is a vital source of employment and economic growth, offering numerous benefits, including environmental, social, and economic advantages. Collaboration among hotels, restaurants, transport providers, and travel agencies enhances the overall efficiency of tourism supply chains and stimulates continuous innovation. Also, the author adds the role of the internet as a transformative communication and marketing tool in the tourism industry, because it enables the creation of platforms such as e-commerce, social media, and mobile applications, reshaping decision-making and traveller behaviour. For more, tourism companies are increasingly investing in digital tools to ensure sustainable success. Information technology (IT) facilitates sustainable practices using social media, chatbots, and digital assistants, promoting best practices and advancing sustainable development goals. These technology-driven strategies play an essential role in meeting tourists' expectations and influencing their decision-making processes. Understanding demographic and cultural factors is also crucial in shaping tourists' perceptions of value (Sanju, 2023).

Information and Communication Technology (ICT) plays a significant role in improving customer experience, enhancing operational efficiency, and supporting strategic decision-making. Key applications of IT/ICT in travel and tourism include:

E-Tourism: Digitalization of processes in tourism, travel, hospitality, and related industries, covering areas such as e-commerce, e-marketing, e-financing, e-accounting, e-HRM, e-procurement, e-strategy, and e-planning.

Specific E-Tourism Applications:

- Flight Tracking System: Monitors and manages global travel activities using aviation software.
- Dynamic Tour Packages: Allows customers to create personalized travel packages in real time.
- Computer Reservation System (CRS): Stores and retrieves information on hotel reservations, airline tickets, and rental cars.
- Global Distribution System (GDS): Facilitates automated transactions between travel service providers.
- Customer Relationship Management (CRM): Manages customer interactions and analyzes data to understand needs.
- Audio Tours: Provides pre-recorded commentary for tourist destinations, used at historical sites and museums.
- GPS Tours: Delivers pre-recorded audio commentary based on visitor location through satellite technology. (Wahab, 2017)

In the tourism industry, IT has become an integral part of business operations, facilitating interaction between companies of all sizes. The adoption of IT requires the redesign of business processes to enhance operational efficiency and adaptability to changing environments. Tourism businesses rely on external information systems for data collection and analysis to gain deeper insights into market dynamics (Labunska, Zyma, & Sushchenko, 2022).

Global distribution and electronic information systems play a vital role in improving the quality and safety of travel services. They provide benefits such as:

- reliability
- adaptability
- communication
- enhancement
- service improvement
- efficient information transfer
- personalized customer experiences

These systems also create effective feedback mechanisms, contributing to the development of the tourism industry (Labunska, Zyma, & Sushchenko, 2022).

IT is essential in the strategic management of tourism businesses, enabling market expansion, cost reduction, employee empowerment, and distribution enhancement. Additionally, the tourism industry is shifting toward cultural and event tourism, utilizing unique attractions to increase regional exposure (Wahab, 2017; Qirici, 2011).

Several European countries have implemented successful tourism information systems, such as the Austria Tyrol and Switzerland Rhodes systems, showcasing how IT can drive growth and improve service quality in the tourism sector (Li, 2010).

2.2 Factors Influencing the Decision-Making Process to Visit a Destination

Numerous studies have explored the factors shaping tourist behaviour and decision-making, highlighting the importance of understanding travel motivations and the elements influencing these decisions. This understanding enables the development of effective promotional systems, helping destinations respond to demand and facilitate the tourism decision-making process.

Climate as a Factor: Hamilton and Lau (2004) emphasized the role of climate, identifying three stages where tourists consider weather conditions: during planning, shortly before the trip, and a week before departure. While seasonal climate remains stable, daily weather variations influence tourists' decisions, leading to a need for weather-based and climate-based information searches (Dahiya & Batra, 2016).

Tourism Decision-Making Variables: Dunne et al. (2011) identified four types of variables:

1. Internal variables: Motivation, attitude, and lifestyle.
2. External variables: Destination attractiveness and social influences.
3. Nature of the trip: Travel time and planned activities.
4. Situational factors: Related to short city break decisions (Dahiya & Batra, 2016).

New Decision-Making Dimensions: Moore et al. (2012) introduced four additional dimensions influencing tourism decisions:

1. Flexibility: Adjusting decisions during the process.
2. Travel location and time: Preferences based on timing and destination.
3. Social composition: Influence from group members like family or friends.
4. Travel stage: Decision-making evolving throughout the journey (Dahiya & Batra, 2016).

Influence of Word of Mouth and social media: In the digital age, online reviews, social media discussions, and recommendations significantly shape tourists' perceptions and decisions. Word of mouth and shared experiences are powerful tools for guiding tourists toward their desired destinations (Dahiya & Batra, 2016).

A study conducted among 861 Romanian tourists identified four key factors influencing destination choice: Destination image (safety, climate, accessibility), Attractions and entertainment (cultural experiences), Service quality (accommodation, price, food), Travel organization (less influential compared to other factors).Together, these factors explained 60.89% of the variance in tourists' decision-making, with destination image emerging as the most critical element (Mihai et al., 2023).

On the other hand, understanding factors for developing an effective information system to promote tourism in Korça is very crucial. By addressing tourists' needs and preferences, this system can streamline decision-making and enhance the travel experience. Below are presented some successful examples.

Czestochowa Tourism Information System (Poland): The Municipal Tourism Information System (MSIT), launched in 2009, illustrates how modern IT can enhance tourism. This system offers comprehensive information about the city's culture, monuments, tourist attractions, and current events, contributing to the city's promotion. The system also provides additional services, including access to local news, employment opportunities, social surveys, and banking transactions. By integrating technology, the system enhances the visitor experience and helps in shaping the city's tourism policies (Biadacz & Biadacz, 2015).

New York Tourism Apps (I Love NY & NYCGO): These digital platforms help attract and manage tourists by curating experiences around iconic landmarks such as Times Square and the Statue of Liberty. They make the city's cultural offerings accessible, while also enhancing tourist experiences by providing information and guiding visitors through the city's attractions. These apps have made New York more competitive in the global tourism market by offering organized, easy-to-navigate itineraries (Dottle, 2016).

Tourism Systems in Hong Kong, Shanghai, and Taipei: Among these three cities, Hong Kong's system stands out for its user-friendliness and comprehensive nature. It offers detailed information on accommodations, attractions, transportation, and customizable itineraries. A digital map highlights key tourist spots, while the system provides audio guides and detailed descriptions of attractions. Shanghai and Taipei also offer valuable tourism information, though Hong Kong's system is noted for its personalized options for diverse traveller needs (Bastida & Huan, 2014).

Austrian Tourism Information System (TIScover): This system emphasizes accessibility through e-commerce and multi-platform access, allowing users to interact via the web, mobile devices, and information kiosks. It integrates high-quality content, enabling local providers to manage their offerings. This personalized approach has proven effective in enhancing Austria's tourism sector (Pröll & Retschitzegger, 2000).

Tourism Information Systems in Morocco: A study of Morocco's tourism systems shows that tourists rely heavily on official tourism websites for information on transportation, accommodation, and attractions. High-quality, detailed content plays a crucial role in shaping tourists' perceptions and decisions, with well-maintained systems positively impacting the country's tourism image (El Maazouzi, 2020).

These case studies highlight the importance of integrating modern information systems in tourism promotion. By offering detailed, accessible information and personalized experiences, cities can attract more tourists and improve their overall visitor experience.

3.Methodology

The research methodology outlined in this thesis adopts a mixed method approach, combining both qualitative and quantitative techniques to develop an information system that enhances user experience and provides essential information to potential visitors of Korça.

The research explores the role of information systems in tourism, examining how technology impacts tourist behaviour, decision-making, and preferences for visiting Korça. The qualitative phase involves reviewing secondary sources such as articles, books, and case studies to analyse the economic importance of tourism, information system components, and the role of technology in tourism promotion. Quantitative research follows, using surveys of 100, 200, and 27 tourists to identify key factors influencing tourist choices, information sources, and desired system features. The results highlight the importance of cultural attractions, recommendations from friends, and the use of websites and social media by tourists.

The research findings guide the development of an information system for Korça's tourism industry, featuring secure data storage and effective information dissemination. Data were collected via secondary sources and online questionnaires, and purposive sampling was used to target tourists who had visited Korça. The analysis combines thematic and content analysis of qualitative data with statistical analysis of quantitative data to identify tourism preferences and inform system design. Ethical considerations were maintained throughout the research, with a focus on integrity and accurate citation of sources.

3.1 Methods and Analysis

To answer the research question and to build an effective system, initially a study with 27 tourists (19 local visitors and 8 foreign tourists) using a questionnaire to assess their experiences, is conducted. The questionnaire focused on identifying their information sources, decision-making factors, desired features in the tourism system, and main issues encountered during their visit. This insight into visitor preferences will guide the development of a promotional system tailored to the needs and desires of tourists in Korça.

Below are presented the results of the study.

1.What sources of information do you typically use to plan your trips?

From the answers result that most respondents relied on recommendations from friends and family for trip planning, with social networks emerging as a key source of information. Some used official tourism websites, while fewer relied on travel apps and online travel agencies.

2.What type of information do you find most helpful when planning your visit to a city?

The study reveals that the history and culture of the city had the greatest influence on destination choice, with 40.7% of respondents citing it as a key factor. City tours followed with 29.6%, and attractions and monuments were noted by 14.8%. Local festivals, shopping, and restaurants had minimal impact, with only 3.7% each for festivals and shopping, and 7.4% for restaurants.

3.How do you prefer to access information while traveling?

From the 26 participants, the majority preferred accessing information through websites or web applications (11 participants). Six preferred social networks, four used mobile applications, three relied on tourist information centres, and two opted for guides and printed brochures.

4.What features would you like to see in an information system for Korça?

Most tourists (69.2%) preferred adding social interactivity to the tourism system, allowing them to share experiences and promote the city. This reflects the importance of recommendations and social networks. Other features, such as detailed advice, online booking, and interactive maps, were favoured by 7.7% each, addressing issues with outdated information and location finding. A small percentage (3.8%) showed interest in augmented reality.

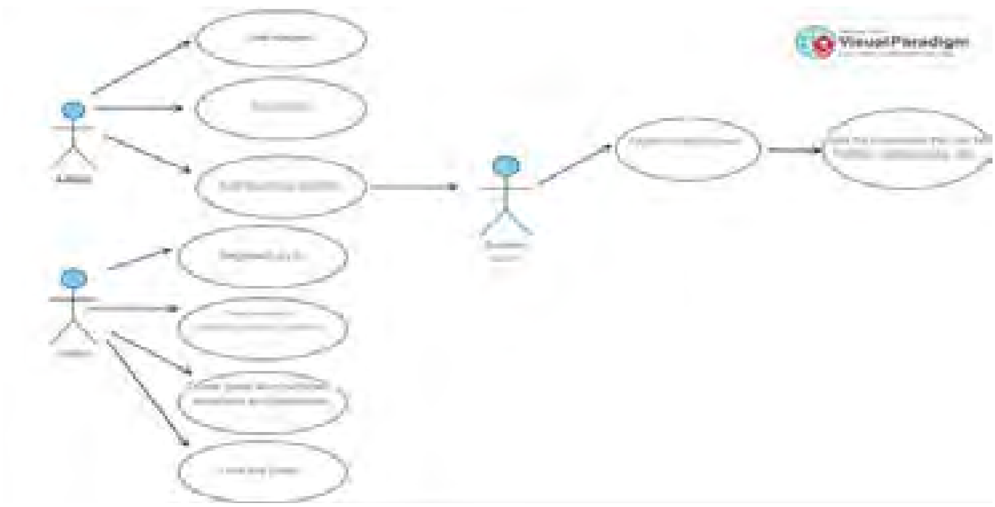
3.2 System's design

The system's design is centred on creating an information system that promotes the natural and cultural attractions, local activities and businesses in Korça. The system targets three user roles: administrators, business owners, and visitors.

- Administrator: Registers business owners and provides them with login credentials via email.
- Business owners: After registration, they can add their businesses to the system for a fee of $10 per listing.
- Visitors: Can explore attractions, activities, and businesses without logging in. However, to post experiences, engage with content, or interact with others, they need to log in and create a profile.

The main goal of the platform is to create a social tourism environment where visitors not only gain valuable insights but also actively participate in promoting the city. By sharing their experiences and engaging with others, visitors become advocates for Korça's tourism, contributing to its growth through social interaction.

FIGURE 1: Use Case Diagram



The technologies used for building the system

For building the described system, we utilized the following technologies:

- Front-End: HTML, CSS, JavaScript
- Back-End: MySQL, Python [Framework: Flask]

Our selection of these technologies was driven not only by personal familiarity and ease of use but also by several objective factors.

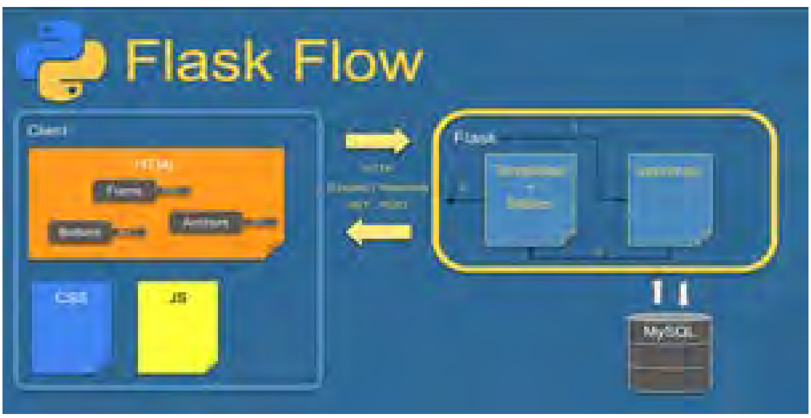
TABLE 1: Advantages (MoldStud, 2024; Coding Dojo, n.d.; Flatirons, 2024; **VYAS, 2023;** Talekar, 2024)

HTML	CSS	Javascript	Python	Flask	MySQL
Wide support	Flexibility and Adaptability	Interactivity	Readability	Modular and minimalist design	Open Source
SEO-friendly	Browser Compatibility	Enhanced User Experience	Libraries and community	Flexibility	Data Security
Enhanced user experience	SEO-friendly	Browser Compatibility	Scope	Extensibility and Integration	Scalability on Demand and higher efficiency
Integration with other technologies	Consistency	Third-Party Integration	Efficacy	Simplicity	Easy to Use
Device compatibility			Integration and Interoperability.		High Performance

How a basic flask application works (flask flow)?
In the image above, we can see how a basic flask application works.
Flow:

1. The HTTP request is made and hits the server.py file.
2. Based on the route we give, it gathers up any HTML, CSS, JS, and data.
3. Then it responds back to the browser with what we return. (Coding Dojo, n.d)

FIGURE 2: Flask Flow (Coding Dojo, n.d)



Database design

The database design for the social-tourism information system focuses on organizing data efficiently to manage user interactions and content. It was developed using MySQL Workbench to facilitate database structuring, including the creation of an Entity Relationship Diagram (ERD) that supports normalized data.

User Roles and Tables:

Administrator: Registers business owners, adds attractions and activities. The Admin table stores essential details (name, email, password, role).

Visitors: Can explore, post, and like content about attractions, activities and businesses. The Visitors table holds their data (name, email, profile picture, etc.).

Business Owners: Add businesses to the system for promotion. The Owners table stores business owner information (name, email, contact details).

Core Functionalities:

Attractions and Activities:

- The Attractions table stores details (name, location, description, etc.).
- The Activities table contains activity data (name, date, time, price, etc.).

Business and Payments:

- The Businesses table includes business-related information (type, description, services, etc.).
- The Payments table tracks payment transactions (amount, date, status).

Social Interactions:

- Visitors can create posts and like content related to attractions, activities, and businesses. Separate tables (PostBusinesses, PostAttractions, PostActivities) store post details, and Likes tables track interactions.
- Testimonials table records user feedback.

Relationships:

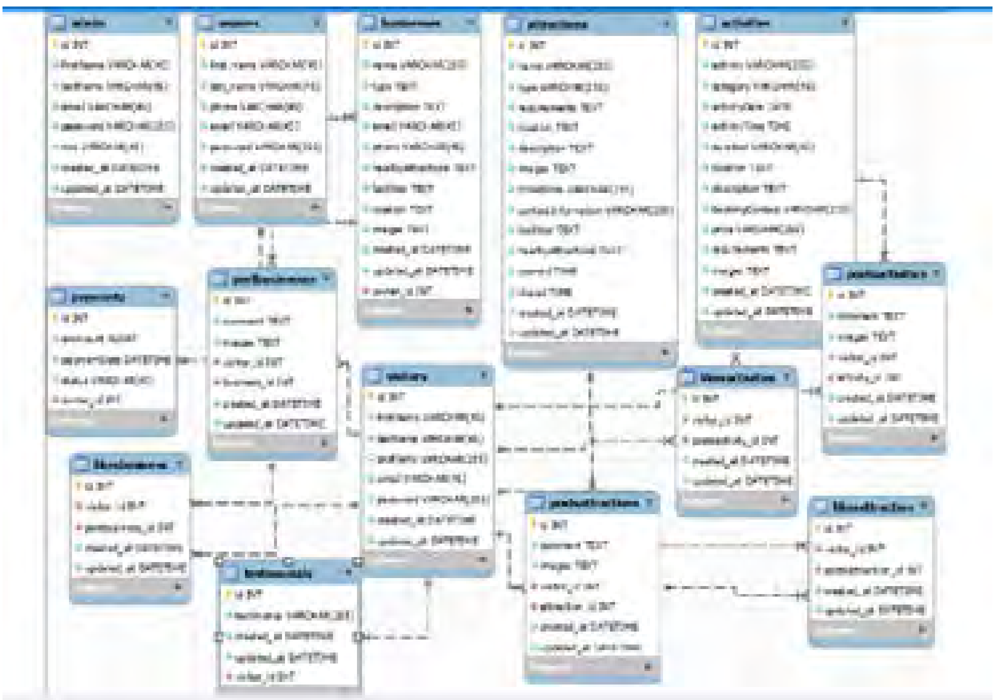
- Owners have a one-to-many relationship with businesses and payments, as each owner can promote multiple businesses and make several payments.
- A visitor can create multiple posts for a specific business, while a business can receive posts from multiple visitors. This establishes a many-to-many relationship between the Visitors and Businesses tables, with the intermediate table PostBusinesses facilitating the connection. The same

structure applies to the relationships between Visitors and Attractions, as well as Visitors and Activities, forming additional many-to-many relationships through PostAttractions and PostActivities tables.Similarly, visitors can like numerous posts, and each post can receive likes from multiple visitors. This creates many-to-many relationships between Visitors and posts related to businesses, attractions, and activities, with the intermediate tables LikesBusiness, LikesAttraction, and LikesActivity managing these interactions.

- Testimonials follow a one-to-many relationship, with each visitor able to leave multiple testimonials.

This structured approach ensures a seamless experience for all users, with each table designed to support the system's functionality efficiently, following best practices in database normalization and relationship management (Letkowski, 2015; Inan & Juita, 2011).

FIGURE 3: ERD diagram



Code structure

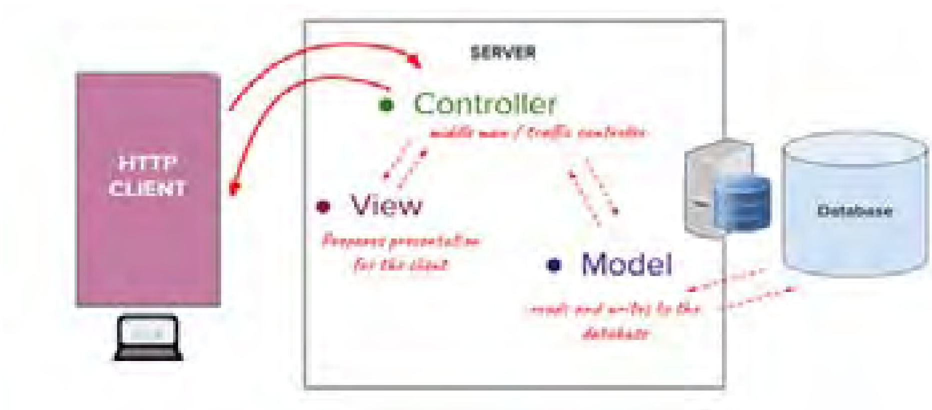
MVC stands for Model, View, and Controller, and is a design pattern used to organize code in a modular way that enhances development, maintainability, and scalability. It is not tied to any specific language or framework.

Controller: The controller manages the initial client request, coordinating between the Model and the View to process and retrieve the necessary data. It acts as a mediator, directing the flow of data between the Model and the View before delivering the final response to the client. It contains the application's core logic.

Model: The model handles data interactions and represents the data entities within the application. It is responsible for reading from and writing to the database. When data is needed, the Model retrieves it and passes it to the Controller, which then forwards it to the View.

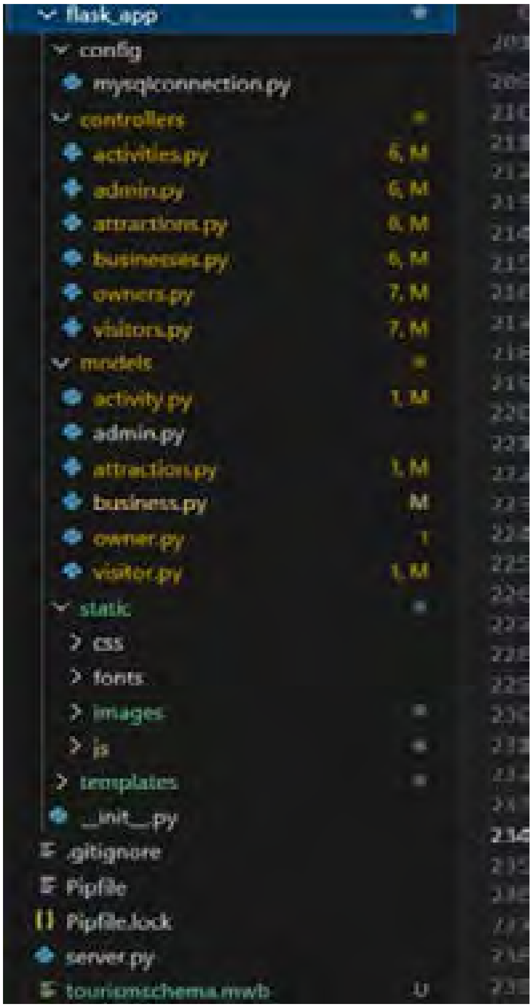
View: The view is responsible for generating the user interface. It takes the data provided by the Controller, integrates it dynamically, and prepares the response in a client-ready format such as HTML. Once the view is rendered, the response is sent back to the Controller, which delivers it to the client. (Coding Dojo, n.d.)

FIGURE 4: MVC design pattern (Coding Dojo, n.d.)



Application modularization
The modular structure of our application is illustrated in the figure below:

FIGURE 5: Application modularization



The modularization approach shown in the figure above organizes our Flask application into distinct components, simplifying management and maintenance.

server.py: This is the primary file of our Flask application, responsible for initializing the app, importing routes, and running the server. Its main role is to configure and execute the Flask application.

The flask_app directory contains:

config folder: Houses configuration files for the application. It includes mysqlconnection.py, which manages the connection to the MySQL database.

controller's folder: Contains functions or routes that handle HTTP requests and return responses.

model's folder: Stores the models representing our application's data structure and facilitating interaction with the database.

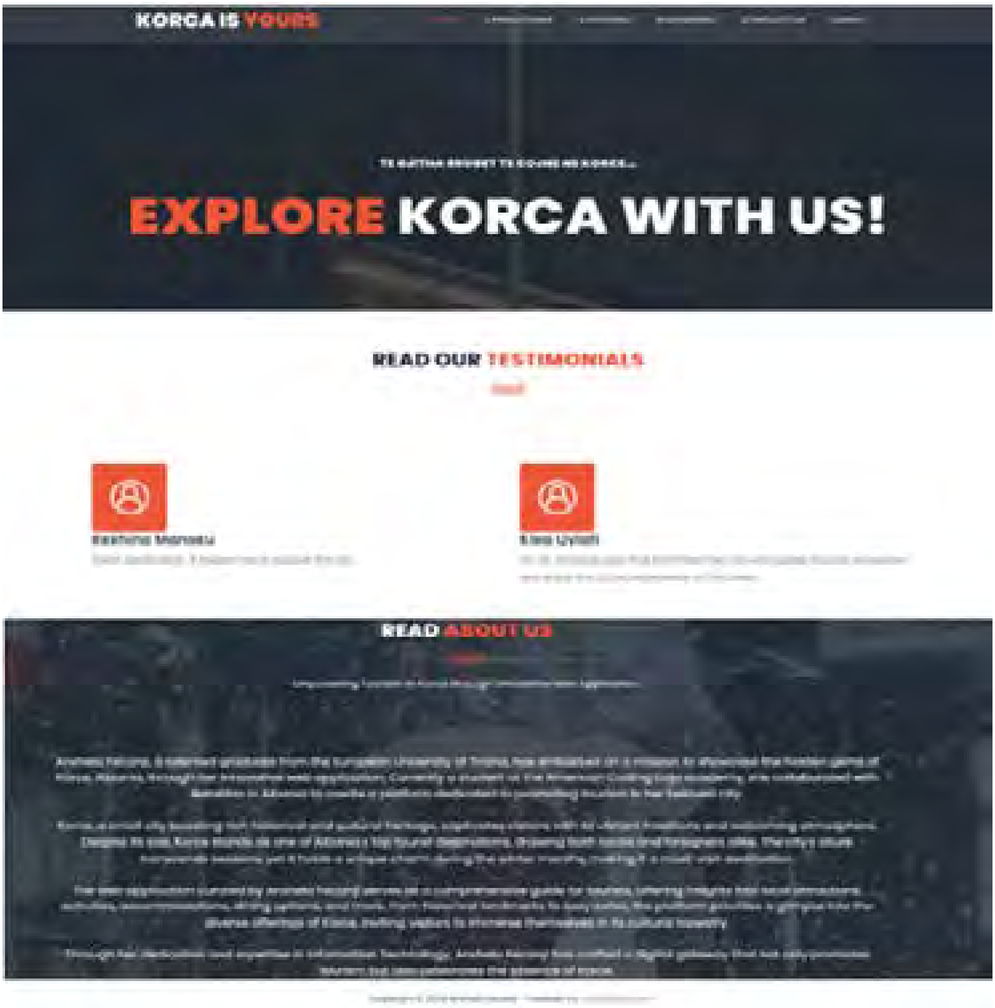
static folder: Holds static resources such as CSS, JavaScript files, and images.

templates folder: Where our HTML pages are stored.

System presentation

Below you can find the presentation of the system.

FIGURE 6: Main Page



Upon entering the main page, visitors will immediately understand its purpose as a promoter of the city of Korça. The page includes information about the developer and the motivation behind the system, along with testimonials from past users, highlighting positive feedback and suggestions for improvements. Visitors can access information about the city's tourist attractions, local businesses such as hotels, restaurants, car rentals, and various cultural activities.

By clicking on "Attractions" in the navigation bar, visitors are directed to a page displaying all the tourist attractions in the city. This page includes a search feature, making it easy to find specific attractions of interest. As shown in the image, each attraction has a "+view more" button. Clicking this button provides additional details, such as the address, description, map location, and photos of the attraction.

FIGURE 7: Attractions presentation

Similarly, clicking on "Activities" in the navbar displays all cultural events in the city, and selecting "Businesses" shows local businesses, each with detailed information.

Clicking "Login" in the navbar opens a dropdown menu with options to log in as either a visitor or an owner.

To log in as an administrator, we will navigate to the route '/loginPage/admin'.

After logging in as an administrator, you are taken to the dashboard, where key functions are accessible via the navigation bar. The administrator can register new business owners, expand the platform's network, and add tourist attractions and city activities. The administrator can add new business owners to the system.

Business owners can contact the site admin via the contact form on the "Contact Us" page. When the admin creates an owner's account, an automated email is sent to the owner with their login credentials.

The administrator can add attractions and cultural activities happening in the city, enhancing the platform's cultural content.

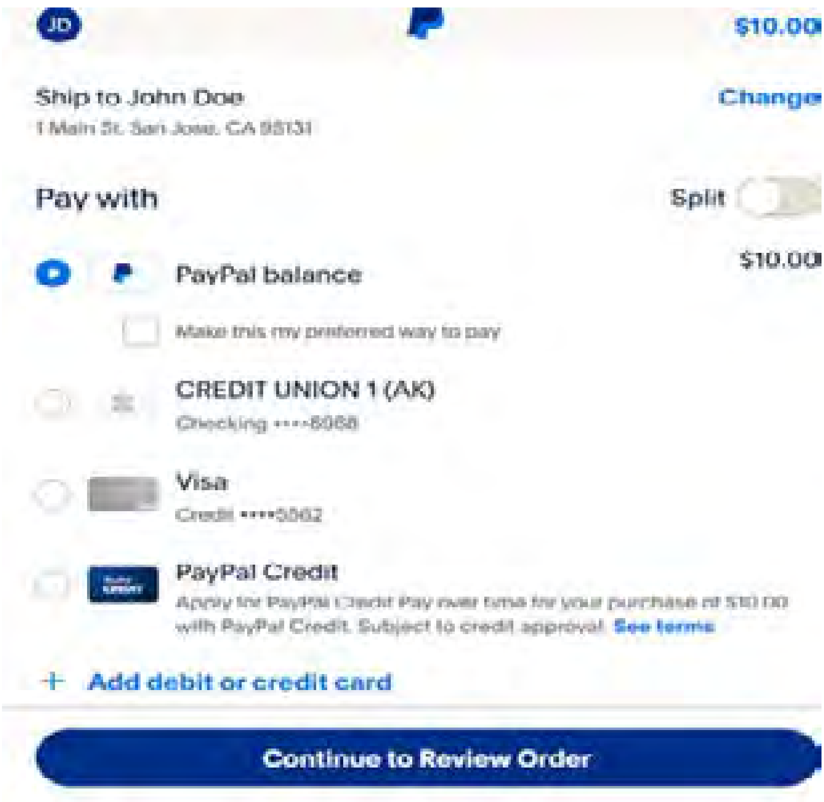
Business Owner

The administrator registers business owners, who can log in on the next page.

The owner dashboard displays all businesses associated with the logged-in owner. Identified owner hasn't added any businesses yet. To register a new business, go to the navbar and click "Add business." You'll be directed to a page to make a standard initial payment set by the administrator before proceeding with the business registration.

Once the payment is completed, a form will appear, allowing us to proceed with the business registration.

FIGURE 8: $10 payment made by owner via PayPal



Visitor

Visitors can log in or register in the application to access detailed information about accommodations, restaurants, tourist attractions, and cultural activities in the city. By joining, they become part of a tourism-social network aimed at promoting the city through visitor-created posts. These posts, based on authentic experiences, act as strong endorsements, highlighting the city's attractions to potential visitors.

After registration, visitors are taken to the dashboard, showing all city businesses, attractions, and activities. Clicking the visitor's profile picture in the navbar redirects to their profile page, which displays posts they've created about businesses, tourist attractions, or cultural activities.

FIGURE 9: Visitor profile page

Clicking the "Delete" button removes the post, while the "View more" button lets registered visitors see users who have liked their post, fostering social interaction between users.

Clicking "All Posts" in the navigation bar opens a page showcasing posts from other visitors. Each post features the visitor's photo, description, creation date and time, and details about the promoted attraction, activity, or business. Clicking on the promoted site provides more detailed information. This highlights the visitor's promotional role. Users can also click "View more" to like the post and see who else has liked it.

3.3 Application security system

In our tourist system, several security measures have been implemented to ensure application security and protect user data:

Data Validation / Input Validation: This process ensures data validity and system integrity by thoroughly validating all user-entered data. This includes validating form submissions during user registration or login, as well as data used for adding new activities, businesses, or attractions.

Authentication and Authorization: These are crucial for verifying user identities and managing access rights to resources and functionality. Users must log in with their unique email addresses and passwords. We enforce strict password security measures, including minimum length requirements and encryption, to protect against unauthorized access (cybertalents, n.d.).

4.Results

This thesis was focused on designing and implementing an effective information system tailored to promoting tourism in Korça, explaining at first, the role of technology in enhancing safety, service quality, and the overall travel experience.

Through the analysis of successful case studies like the I Love NY and NYCGO apps, as well as systems in Hong Kong, Shanghai, Taipei, and Morocco, is shown how such systems can boost the tourism industry and improve visitor experiences. This study highlighted key factors influencing destination choices, such as accessibility, attractiveness, affordability, and amenities.

Focusing on Korça, are identified natural and cultural attractions as primary tourist draws, with online platforms and social interaction features being highly desired by tourists. Recognizing Korça's significance as a major Albanian tourist destination, was developed a system that provides detailed information on attractions, accommodations, and dining, while facilitating engagement between businesses and visitors.

This thesis provides a guide for developing an information system tailored to promote tourism in Korça. The system aims to create a relaxed, social, and engaging experience for tourists, stimulate economic growth, and support sustainable tourism practices. By bringing together stakeholders and visitors, the system seeks to highlight Korça's diverse offerings and enhance its status as a major tourist destination.

Future work

The future work aims:

- The transformation into a dynamic social-tourist network dedicated to Korça
- Organization of group city tours through the system
- Online booking option for accommodation and restaurants
- Multilingual support
- Personalized recommendations

References:

1. Bastida, U., & Huan, T. (2014). Performance evaluation of tourism websites' information quality of four global destination brands: Beijing, Hong Kong, Shanghai, and Taipei. Journal of Business Research,67(2), 167–170.
2. Biadacz, R., & Biadacz, M. (2015). The Use of Modern Information Technology in Tourist Information Systems on the Example of City of Czestochowa. Procedia Computer Science,65, 1105–1113.
3. Coding Dojo. (n.d.). Flask-Bcrypt. Retrieved from: https://login.codingdojo.com/m/506/12465/87444
4. Coding Dojo. (n.d.). MVC Design Pattern. Retrieved from: https://login.codingdojo.com/m/506/12464/110535
5. Coding Dojo. (n.d.). Overview. Retrieved from: https://login.codingdojo.com/m/506/12467/87456
6. Coding Dojo. (n.d.). Python Overview. Retrieved from: https://login.codingdojo.com/m/506/12456/87290
7. Dahiya, Kirti & Batra, Dr. (2016). Tourist decision making: Exploring the destination choice criteria. Asian Journal of Management Research. 7. 140-153. Retrieved from: https://www.researchgate.net/publication/325192595_Tourist_decision_making_Exploring_the_destination_choice_criteria
8. Dino, A. (2018). Turizmi kulturor si mundësi për një zhvillim të qëndrueshëm në rajonin e Korçës (punim doktorature, Universiteti Europian i Tiranës, Tiranë, Shqipëri). Retrieved from: https://uet.edu.al/wp-content/uploads/2019/07/Doktoratura-AVD-2018-1.pdf
9. Dottle, R. (2016, January 30). Geographies of the Tourism Economy, Space Way-finding, New York and Spectacle. Columbia. Retrieved from: https://www.academia.edu/21244696/Geographies_of_the_Tourism_Economy_Space_Way_finding_New_York_and_Spectacle
10. Dunne, G., Flanagan, S., & Buckley, J. (2011), Towards a decision-making model for city break travel. International Journal of Culture, Tourism and Hospitality Research, 5(2), pp 158-172.
11. El Maazouzi, A. (2020). The Impact of Official Tourism Websites on the Destination Image: The case of Morocco (Dissertation). Retrieved from: https://www.diva-portal.org/smash/get/diva2:1474902/FULLTEXT01.pdf
12. Flatirons. (2024, January 03). What is Flask? Overview of the Flask Python Framework. flatirons.com. Retrieved from: https://flatirons.com/blog/what-is-flask-overview-of-the-flask-python-framework-2024/
13. INSTAT Albania. (2020). Turizmi në shifra: Shqipëri 2020. Retrieved from: https://www.instat.gov.al/media/9545/turizmi-ne-shifra-shqiperi-2020__.pdf
14. INSTAT Albania. (2021). Turizmi në shifra: Shqipëri 2021. Marrë nga https://www.instat.gov.al/media/11106/turizmi-ne-shifra-shqiperi-2021.pdf
15. Korça, Qyteti im. Turist në Korçë. (2016). https://turistnekorce.wordpress.com/home/
16. Labunska, S., Zyma, O., & Sushchenko, S. (2022). The use of information systems as a way to ensure interaction between small and big tourism enterprises. Access Journal - Access to Science, Business, Innovation in the Digital Economy,3(1), 16–28. https://doi.org/10.46656/access.2022.3.1(2)
17. Lau, J. M. (2004), The role of climate information in tourist destination choice. Hamburg, Germany.

18. Li, H. (2010). Study on Tourism Planning and Management Information System Based on Component GIS Technology. Proceedings of Annual Conference of China Institute of Communications. Retrieved from: https://www.scirp.org/pdf/6-1.41.pdf
19. Mihai, V. C., Dumitras, D. E., Oroian, C., Chiciudean, G. O., Arion, F. H., & Mureșan, I. C. (2023). Exploring the Factors Involved in Tourists' Decision-Making and Determinants of Length of Stay Administrative Sciences, 13(10), 215.
https://doi.org/10.3390/admsci13100215
20. Ministria e Turizmit dhe Mjedisit e Shqipërisë. (2019). Strategjia Kombëtare e Turizmit 2019-2023.Retrieved from://turizmi.gov.al/wp-content/uploads/2019/06/Strategjia-Komb%C3%ABtare-e-Turizmit-2019-2023.pdf
21. MoldStud. (2024). The Importance of HTML, CSS, and Javascript in Web Development. Retrieved from: https://moldstud.com/articles/p-the-importance-of-html-css-and-javascript-in-web-development
22. Moore, K., Smallman, C., Wilson, J., & Simmons, D. (2012), Dynamic in-destination decision-making: An adjustment model. Tourism management, pp 635-645.
23. Partners Albania. (2015). Promovimi_i_turizmit_alternativ_Qarku_Korce. Partners Albania, Center for Change and Conflict Management. Retrieved from: http://partnersalbania.org/wp-content/uploads/2015/10/Broshure_Promovimi_i_turizmit_alternativ_Qarku_Korce.pdf
24. Pröll, B., & Retschitzegger, W. (2000). Discovering Next Generation Tourism Information Systems: A Tour on TIScover.Journal of Travel Research,39(2), 182-191. https://doi.org/10.1177/004728750003900208
25. Qirici, E. (2011). Analysis of Consumers Profile as an Important Tool for Tourism Development. Journal of Knowledge Management, Economics and Information Technology. Retrieved from: https://ideas.repec.org/a/spp/jkmeit/1222.html
26. Qirici, E. (2014). Events as a Source of Competitive Advantage for Destinations. European Scientific Journal, 18(0). Retrieved from:
https://eujournal.org/index.php/esj/article/view/4681/4473
27. Sanju, N. L. (2023). Trends of Information Systems in Tourism: A Review of Literature. International Journal of Research and Review,10(12), 229–241. https://doi.org/10.52403/ijrr.20231227
28. Talekar, S. (2024). 6 Reasons to Choose Python for Backend Development. Retrieved from: https://www.thinkitive.com/blog/6-reasons-to-choose-python-for-backend-develop
29. Vyas, K. (2023). 8 Big Advantages of Using MySQL | Datamation. Retrieved from: https://www.datamation.com/storage/8-major-advantages-of-using-mysql/
30. Wahab, I. (2017). Role of Information Technology in Tourism Industry: Impact and Growth. International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE). Retrieved from: https://www.researchgate.net/publication/363040658_Role_of_Information_Technology_inTourism_Industry_Impact_and_Growth
31. Xhajanka, E. (2021). Korça qyteti i preferuar i turistëve të huaj [Korça, the preferred city of foreign tourists]. Retrieved from: https://ata.gov.al/2021/08/01/korca-qyteti-i-preferuar-i-turisteve-te-huaj/
32. Zaimi, E. (2017). Korça, turizëm në katër stinët. Revista Monitor. Retrieved from: https://www.monitor.al/korca-turizem-ne-kater-stinet-3/
33. World Tourism Organization. (2024). World Tourism Barometer: January 2024. Retrieved from: https://en.unwto-ap.org/resources/worldtourismbarometer-january-2024/
34. UNWTO. (2024). International Tourism to Reach Pre-Pandemic Levels in 2024. Retrieved from: https://www.unwto.org/news/international-tourism-to-reach-pre-pandemic-levels-in-2024

The Use of Artificial Intelligence in Building Sign Language Recognition Application

Ditmira Tahiri[1]
European University of Tirana

Ardiana Topi
European University of Tirana

Abstract

Communication serves as the foundation upon which societies are built, manifesting in various forms such as gestures, sounds, drawing, writing, and speech. However, for individuals with hearing impairments, traditional modes of communication like spoken language can pose substantial challenges. These challenges often result in barriers to effective interaction, not only in personal and social settings but also in educational and professional environments. To bridge this gap, sign language has emerged as an essential and empowering communication tool, enabling individuals with hearing impairments to express themselves with clarity and nuance. Sign language is not just a series of gestures but a fully developed language system. It serves as a vital channel through which individuals with hearing impairments can interact with the world around them, breaking down the barriers that their condition imposes.

[1] Ditmira Tahiri graduated with a degree in *Computer Engineering* from *Trakya University* in Turkey. During the final year of her studies, she gained her first professional experience as a *test engineer* at the German company named Ronwell, in Turkey, where she worked for nearly one and a half years. After returning to Albania, she pursued a *Master of Science in Software Engineering* at the *European University of Tirana (UET)*. She briefly worked at a programming company in Shkodra before transitioning to the role of *IT specialist* in the IT department at the Municipality of Shkodra.

Despite technological advances, the recognition and interpretation of sign language remain complex tasks, especially given the inherent complexity of sign languages characterized by multiple channels, including manual gestures, facial expressions, and body language.

This study leverages the Python programming language and the YOLOv8 object detection framework to develop a practical sign language recognition application. This system utilizes deep learning and computer vision to interpret sign language gestures in real-time, aiming to address the limitations of existing recognition systems.

The developed system achieved a 95% accuracy rate in recognizing sign language gestures, demonstrating the effectiveness of combining Python with YOLOv8 for this application.

This research contributes to the field of assistive technologies by providing a

versatile and user-friendly tool that can be deployed across various platforms and environments, ultimately enhancing communication and social integration for individuals with hearing impairments.

Keywords: *Sign Language Recognition, Deep Learning, YOLOv8, Python, Assistive Technology.*

1. Introduction

1.1. Background and Importance of Sign Language

Communication is an essential pillar of human interaction and societal development, manifesting in various forms such as spoken language, written text, visual symbols, and gestures. It enables individuals to share ideas, express emotions, and convey complex thoughts, playing a critical role in the cohesion and functioning of societies. For most people, spoken language serves as the primary means of communication, providing an efficient and direct way to engage with others. However, for individuals who have hearing impairments or hard of hearing, traditional modes of verbal communication present significant challenges. The inability to hear and respond to auditory cues can lead to feelings of isolation, frustration, and exclusion from mainstream social interactions. This can have profound implications, not only for personal and social development but also in accessing essential services such as education, healthcare, and employment opportunities. In such cases, sign language emerges as a crucial tool for bridging this communication gap.

Sign language is much more than a series of hand gestures; it is a fully developed and complex language system with its own grammar, syntax, and vocabulary. It allows individuals with hearing impairments to express themselves with precision and nuance, using a combination of hand movements, facial expressions, and body language to convey messages effectively. Used by millions of people worldwide, sign language is an integral part of the community of people with hearing loss, fostering a sense of identity and belonging. However, despite its importance, sign language users often face significant challenges in communication, especially in environments where interpreters are not available or where there is a lack of understanding and support for their language. This can lead to social exclusion and limit their access to various public services, educational resources, and employment opportunities. The scarcity of interpreters and the high costs associated with their services further exacerbate these challenges, making it difficult for individuals who rely on ,sign language to participate fully in society.

The development of technology that can bridge the communication gap for sign language users is therefore of paramount importance. Technological advancements, particularly in the field of artificial intelligence (AI), offer promising solutions to these challenges by enabling the creation of systems that can recognize and interpret sign language in real-time. Such systems can serve as a vital resource for enhancing communication and interaction, thereby promoting greater inclusion and accessibility for individuals with hearing impairments. The ability to use technology to translate sign language into spoken or written text instantly can significantly improve the quality of life for these individuals, providing them with more opportunities to engage with the world around them. It is essential to continue developing and refining these technologies to ensure they are accurate, user-friendly, and accessible to all who need them, ultimately helping to create a more inclusive and equitable society.

1.2. The Role of Artificial Intelligence in Assistive Technologies

Artificial Intelligence (AI) has rapidly evolved over recent years, becoming a transformative tool across multiple fields, including healthcare, finance, and education. In the realm of assistive technologies, AI has demonstrated remarkable potential in improving the quality of life for individuals with disabilities by offering innovative solutions to complex problems. AI systems, especially those utilizing machine learning and computer vision, are capable of performing tasks that were once considered to require human intelligence. These systems can analyze vast amounts of data, recognize patterns, and make decisions based on the information they process, enabling them to carry out a wide range of functions, from recognizing speech and text to interpreting visual cues. One of the most promising applications of AI in assistive technology is in the field of sign language recognition. AI-driven sign language recognition systems are designed to bridge the communication gap between sign language users and those who do not understand the language. These systems use advanced machine learning algorithms and deep learning models to process visual data, such as images or video sequences of sign language gestures, and translate them into spoken or written language in real-time. This technology provides a powerful tool for communication, especially in situations where human interpreters are not available. By leveraging the capabilities of AI, these systems can significantly reduce the reliance on human interpreters, offering an always-available, cost-effective solution for individuals who use sign language. Moreover, AI-driven systems have the capacity to learn and improve over time. They can be trained on diverse datasets to recognize various sign languages and dialects, making them adaptable to different linguistic contexts and enhancing their utility for a global user base.

The integration of AI into sign language recognition systems presents numerous advantages. First, it democratizes access to communication tools for individuals with hearing impairments, providing them with the means to interact more freely and effectively in diverse settings. This can lead to improved outcomes in education, employment, and social integration, as users can engage more fully in conversations, participate in classroom discussions, and access services without the constant need for an interpreter. Additionally, AI-driven sign language recognition systems can be incorporated into various digital platforms, such as smartphones and computers, making them accessible to users wherever they are. This portability and convenience are crucial for enhancing the everyday experiences of sign language users, enabling them to communicate more fluidly in both personal and professional contexts. Furthermore, the adaptability of AI technologies allows these systems to be continuously refined and updated, improving their accuracy and efficiency in recognizing and translating sign language. As more data becomes available and as machine learning models become more sophisticated, AI-driven systems will be able to recognize an even broader range of gestures and facial expressions, further enhancing their effectiveness. This ongoing development is essential for ensuring that these technologies remain relevant and useful as the needs of the community of people with hearing loss evolve. In conclusion, the integration of AI into assistive technologies, particularly in sign language recognition, holds significant promise for empowering individuals with hearing impairments and fostering a more inclusive society where communication barriers are minimized.

1.3. Study Objective and Research Questions

The primary objective of this study is to develop a sign language recognition system that utilizes AI to facilitate seamless communication for individuals with hearing impairments. The system aims to achieve high accuracy in real-time sign language interpretation by using deep learning techniques, specifically the YOLOv8 object detection framework, in combination with Python. The study employs the YOLOv8 (You Only Look Once) object detection framework, which is known for its efficiency and precision in processing visual data, combined with the Python programming language, which offers robust libraries and tools for implementing complex machine learning models. The objective is not only to build a functional prototype but also to push the boundaries of current sign language recognition capabilities, making the system both practical and adaptable to various real-world scenarios.

This research seeks to address the following questions:

- What is the impact of AI on the accuracy and efficiency of sign language recognition?
- How can a sign language recognition system be implemented using Python and deep learning frameworks?
- What are the challenges and limitations associated with current sign language recognition technologies?
- How can the developed system be integrated into educational settings to enhance learning experiences for students with hearing impairments?

2. Literature Review

2.1. Evolution of Artificial Intelligence in Assistive Technologies

Artificial Intelligence (AI) has experienced a profound transformation over the past few decades, evolving from a collection of theoretical concepts and speculative possibilities into a powerful and versatile tool with practical applications across numerous fields. This evolution has been particularly impactful in the development of assistive technologies, where AI has emerged as a cornerstone in creating innovative systems designed to enhance the quality of life for individuals with disabilities. From speech recognition software that enables individuals with speech impairments to communicate more effectively to AI-driven prosthetics that offer improved mobility and dexterity for amputees, the influence of AI on assistive technologies is both broad and deep. These advancements have not only expanded the functionality of assistive devices but also significantly broadened the horizons of what is possible, empowering individuals with disabilities to engage more fully in educational, professional, and social environments.

2.2. Sign Language Recognition Systems: A Historical Perspective

The development of sign language recognition systems has followed a trajectory similar to that of other AI applications, beginning with rudimentary models and evolving into sophisticated, data-driven systems capable of handling complex tasks. Early efforts in this field were primarily focused on developing algorithms that could recognize static hand signs. These initial systems employed basic image processing techniques to detect the shape and position of the hand, attempting to map these features to specific signs. However, they faced significant limitations in their ability to function reliably in diverse settings. Variations in lighting, background, and hand orientation often led to inaccuracies, making these early systems impractical for real-world use. Additionally, these systems were typically limited to recognizing a small set of predefined signs, which greatly constrained their utility in real-world communication scenarios where a much larger vocabulary of signs is used.

As technology advanced, researchers began to explore the potential of machine learning to enhance the capabilities of sign language recognition systems. Unlike rule-based systems that required explicit programming of every possible scenario, machine learning algorithms could learn from large datasets of sign language gestures, enabling them to recognize a much wider range of signs with greater accuracy and robustness. This shift from rule-based to data-driven approaches marked a significant milestone in the field. Neural networks, particularly Convolutional Neural Networks (CNNs), played a crucial role in this transformation. These models could automatically learn to identify and extract relevant features from raw image data, such as hand shapes and movements, without the need for manual feature engineering. This capability allowed researchers to develop more generalizable models that could handle a variety of sign languages and dialects.

The use of machine learning and deep learning in sign language recognition has continued to evolve, with modern systems now capable of recognizing not only static hand signs but also dynamic sequences of gestures. This has significantly expanded the range of applications for these systems, making them useful not only for basic communication tasks but also for more complex interactions that involve continuous sign language. Despite these advancements, challenges remain, particularly in achieving high levels of accuracy across different sign languages and in varying environmental conditions. The historical development of sign language recognition systems thus reflects a broader trend in AI research: the movement from simple, narrowly focused models to more sophisticated, flexible systems capable of addressing complex, real-world problems.

2.3. Deep Learning in Sign Language Recognition

Deep learning, a subset of machine learning, has revolutionized the field of computer vision and has had a particularly profound impact on the development of sign language recognition systems. Deep learning models, especially

Convolutional Neural Networks (CNNs), have proven to be highly effective in tasks involving image recognition due to their ability to automatically learn hierarchical representations of data. Unlike traditional machine learning algorithms that require manual feature extraction, CNNs can directly process raw pixel data, learning to identify patterns and features such as edges, shapes, and textures through multiple layers of abstraction. This makes them exceptionally well-suited for recognizing complex visual patterns, such as the intricate hand shapes, movements, and facial expressions involved in sign language.

In the context of sign language recognition, CNNs have been used to develop systems that can accurately identify and interpret a wide range of gestures. These models are trained on large datasets of sign language videos, learning to recognize the subtle nuances of different hand shapes and movements. The use of deep learning has enabled significant improvements in the accuracy and robustness of sign language recognition systems, allowing them to handle a greater variety of signs and to operate effectively in diverse environmental conditions. One of the most advanced frameworks in this domain is the YOLO (You Only Look Once) object detection model, which is known for its ability to detect and classify objects in images quickly and accurately. The latest iteration of this framework, YOLOv8, offers enhanced performance, making it ideal for real-time sign language recognition applications.

By using YOLOv8 in combination with Python, this study aims to create a system that can accurately and efficiently recognize sign language gestures in real-time. The real-time aspect is particularly important for applications that require immediate feedback, such as communication aids for individuals with hearing impairments or educational tools for teaching sign language. The integration of deep learning techniques into sign language recognition systems represents a significant advancement, enabling these systems to achieve levels of performance that were previously unattainable. However, there are still many challenges to be addressed, particularly in terms of improving the generalizability of these models across different sign languages and dialects, as well as enhancing their ability to operate in complex, real-world environments.

2.4. Challenges in Sign Language Recognition

Despite the significant advancements made possible by AI and deep learning, sign language recognition remains a highly challenging task. One of the primary challenges is the inherent diversity of sign languages and their dialects. Unlike spoken languages, which often have standardized written forms and well-defined grammatical rules, sign languages can vary significantly from one region to another, and even within the same community. This diversity poses a substantial challenge for AI systems, which must be trained on large and diverse datasets to achieve high accuracy. The collection of such datasets is often difficult due to the lack of standardized sign language resources and the variability in sign language use across different populations.

Another major challenge is the dynamic nature of sign language itself. Unlike static hand signs, which can be recognized using traditional image processing techniques, continuous sign language involves complex sequences of movements that must be interpreted in context. This requires AI systems not only to recognize individual gestures but also to understand the temporal and spatial relationships between them. The ability to accurately interpret these dynamic sequences is crucial for achieving high levels of recognition accuracy, but it adds a layer of complexity that is difficult to manage. Deep learning models, such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, have been used to address these challenges by modeling the temporal dependencies between gestures, but these models often require large amounts of data and significant computational resources. Finally, the integration of sign language recognition systems into real-world applications presents its own set of challenges. For these systems to be practical and useful, they must be robust enough to handle variations in lighting, background, and camera angles, and they must be able to operate in real-time. Achieving this level of robustness and efficiency requires careful tuning of the model parameters and extensive testing in diverse environments. Additionally, there are challenges related to the usability and accessibility of these systems, as they must be designed in a way that is intuitive and easy for users to operate. Overcoming these challenges will be essential for the widespread adoption of sign language recognition technologies, and it will require ongoing research and development to refine these systems and ensure that they meet the needs of all users.

3. Methodology

3.1. Theoretical Framework

The theoretical framework for this study is based on the principles of deep learning and computer vision. Deep learning, particularly Convolutional Neural Networks (CNNs), has been shown to be highly effective in image recognition tasks. CNNs are capable of learning hierarchical representations of data, which makes them well-suited for recognizing complex patterns in images, such as those found in sign language gestures.

The YOLO (You Only Look Once) framework, which is used in this study, is a single-stage object detection model that has gained popularity due to its speed and accuracy. YOLOv8, the latest version of this framework, offers several improvements over previous versions, including better handling of small objects and more accurate bounding box predictions. By combining YOLOv8 with Python, this study aims to create a system that can recognize sign language gestures in real-time with high accuracy.

3.2. Practical Framework

The practical framework for this study centers on the development of a sign language recognition system utilizing Python and the YOLOv8 object detection framework. This system is designed to recognize a predefined set of sign language gestures and translate them into corresponding text in real-time. The overarching goal of this framework is to provide a highly accurate, user-friendly tool that can be used by individuals who rely on sign language for communication, facilitating smoother interaction in various settings. The development process is composed of several critical stages, each contributing to the overall functionality and effectiveness of the system. These stages include data collection, model training, system implementation, and thorough testing and evaluation

Data Collection: The first step in the development process is data collection, which is essential for training the sign language recognition model. In this stage, a diverse and comprehensive dataset of sign language gestures is gathered, encompassing various hand shapes, movements, and facial expressions. To ensure the system's versatility and accuracy, the dataset must represent a wide range of sign languages and dialects, reflecting the diversity within the hard-of-hearing communities. This diversity allows the system to recognize and interpret gestures more effectively, minimizing bias and ensuring it is applicable in different linguistic and cultural contexts. The collected dataset forms the foundation upon which the entire recognition system is built, providing the raw data necessary for training the model to accurately detect and interpret sign language gestures.

Model Training: Once the dataset is collected, the next step is model training. This is a critical step, as it determines the system's ability to recognize and interpret gestures accurately. The YOLOv8 model, known for its efficiency and precision in object detection, is trained using the labeled dataset of sign language gestures. During training, the model is fed with data that pairs specific gestures with their corresponding translations into text. The supervised learning approach is employed here, which allows the model to adjust its parameters and minimize prediction errors based on feedback from the training data. Through this iterative process, the model improves its accuracy and learns to generalize better across different gestures, ensuring that it can recognize a wide range of sign language inputs.

System Implementation: After the model is successfully trained, the next crucial step is its integration into a fully operational real-time sign language recognition system. This system is specifically designed to take live input from a camera, which captures video frames containing the gestures made by a user. Each frame is then processed using the trained YOLOv8 model, which is highly efficient at object detection and well-suited for recognizing the complex hand movements and gestures that constitute sign language. The YOLOv8 model identifies and classifies these gestures in real-time, detecting not only the hand shapes but also the dynamic motions and any accompanying facial expressions that may add meaning to the gestures. Once the model recognizes the gestures, it translates them into corresponding text, providing an immediate, readable output for the user or any interacting systems. The system's implementation is achieved using Python, a versatile programming language that supports powerful libraries necessary for complex tasks like image and video processing. One of the key tools utilized in this process is OpenCV, an open-source computer vision library that enables efficient processing of video data. OpenCV allows the system to capture and handle video frames from the camera in real-time, seamlessly feeding them into the YOLOv8 model for analysis. This integration of the YOLOv8 framework and OpenCV within Python creates a robust and efficient pipeline that enables the real-time recognition of sign language gestures. The system's design ensures that it operates fluidly and quickly, providing instant feedback by converting visual inputs into text. This makes the tool highly practical for real-world applications, where immediate translation and communication are critical, particularly in educational and assistive contexts where sign language plays a central role.

Testing and Evaluation: The final step is to test and evaluate the performance of the sign language recognition system. The system is tested on a separate dataset of sign language gestures that were not used during training. The accuracy of the system is measured by comparing its predictions with the ground truth labels. The system's performance is also evaluated in real-world scenarios, such as in educational settings, to assess its practicality and robustness.

3.3. Development of the Sign Language Recognition System

The development of the sign language recognition system requires the execution of several technical steps, each of which is fundamental to ensuring that the system can accurately and efficiently recognize sign language gestures in real-time, including data preprocessing, model training, and system integration.

The first key step is **data preprocessing.** Before the model can be trained, the raw data must be preprocessed in a format that can be effectively used by the model, to ensure that it is suitable for training. This involves tasks such as resizing the images, normalizing the pixel values, and augmenting the data to increase the diversity of the training set. Data augmentation techniques, such as rotation, flipping, and scaling, are often employed at this stage to expand the training dataset and improve the model's ability to generalize across diverse environments and also improve the robustness of the model.

Once the data has been preprocessed, the next critical step is model **training.** In this phase, the YOLOv8 model is trained on the preprocessed dataset using a supervised learning approach.

The training process involves feeding the model with labeled examples of sign language gestures and adjusting its parameters to minimize the error in its predictions. The model is trained using backpropagation, a standard algorithm in deep learning that updates the model's parameters based on the gradient of the loss function.

Following the training process, the final step is **system integration**, where the trained YOLOv8 model is embedded into a fully functional sign language recognition system. This involves creating a seamless pipeline that connects the input from a camera, processes the video frames using the trained model, and outputs the recognized gestures as text in real-time. The integration process is handled using Python, leveraging the YOLOv8 framework for object detection and OpenCV for video processing. The combination of YOLOv8 and OpenCV allows the system to operate with remarkable speed and accuracy, ensuring that sign language gestures are recognized and translated into text almost instantly. This real-time capability is essential for the system's practical applications, making it suitable for use in communication, education, and assistive technology contexts where immediate feedback is necessary.

4. Methods and Analysis

4.1. System Architecture

The architecture of the sign language recognition system is based on the YOLOv8 object detection model, which is a single-stage detector that can recognize multiple objects in an image in real-time. The model takes an input image, divides it into a grid, and predicts bounding boxes and class probabilities for each grid cell. The YOLOv8 model is known for its speed and accuracy, making it well-suited for real-time applications.

In the context of sign language recognition, the input to the YOLOv8 model is a video frame captured by a camera. The model processes the frame and outputs a set of bounding boxes, each representing a detected hand or face. The model also predicts the class of each detected object, which corresponds to a specific sign language gesture.

The output of the YOLOv8 model is then processed to generate the final translation. This involves combining the detected gestures into a coherent sequence and mapping them to the corresponding words or phrases in the target language. The translation is displayed as text on the screen, providing real-time feedback to the user.

4.2. Pose Evaluation

Pose evaluation is a critical and complex component of the sign language recognition system, as it requires the system to accurately analyze the position, movement, and orientation of the hands and face to determine the meaning of the gestures being performed. In sign language, both hand shapes and facial expressions play an integral role in conveying nuanced meanings, making it essential for the system to interpret these elements with precision. This task typically involves a combination of computer vision techniques and deep learning models, which work together to track and evaluate the gestures in real-time. The system must be able to not only detect the hands and face but also recognize the intricate movements and changes in expression that occur during communication.

In this study, pose evaluation is carried out using the YOLOv8 model, a highly advanced object detection framework. YOLOv8 is used to detect the hands and face in each input image or video frame, predicting their precise positions and movements. Once the relevant features are detected, the model classifies the objects into various categories, such as specific hand shapes or facial expressions. This classification is crucial for understanding the meaning behind the gestures, as different hand configurations and facial cues can represent different words, phrases, or grammatical elements in sign language. The pose information extracted from the YOLOv8 model is then processed to interpret the sign language gestures, enabling the system to generate the corresponding translation in text form.

Pose evaluation presents significant challenges in the context of sign language recognition due to the wide variability in hand shapes, movements, and facial expressions across different signers and situations. For example, the same gesture may be performed differently depending on the individual's signing style, the speed of their movements, or the context in which the gesture is used. To overcome these challenges, the YOLOv8 model is trained on a large and diverse dataset that includes a wide range of sign language gestures. This dataset allows the model to learn the natural variations in signing and adapt to different contexts, improving its ability to accurately recognize gestures across a broad spectrum of users and environments. By training the model on diverse data, the system becomes more robust and effective at handling the inherent complexity of sign language, ensuring that it can provide accurate and reliable translations in real-time.

4.3. Sign Language Recognition using Sensors and Machine Learning

In addition to computer vision techniques, sign language recognition can also be performed using sensors and machine learning models. Sensors, such as accelerometers and gyroscopes, can be used to capture the motion of the hands and body, providing additional information for gesture recognition.

Machine learning models, such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, can be used to process the sensor data and recognize the sign language gestures. These models are particularly effective for recognizing sequences of gestures, as they can capture the temporal dependencies between the gestures.

In this study, the YOLOv8 model is combined with machine learning techniques to improve the accuracy and robustness of the sign language recognition system. The model is trained on a dataset that includes both visual and sensor data, allowing it to leverage the strengths of both approaches and achieve better performance.

5. Results

5.1. Accuracy of the Recognition System

The sign language recognition system developed in this study achieved an impressive accuracy rate of 95% in identifying sign language gestures. This level of success was determined through rigorous testing on a separate dataset that had not been utilized during the training phase. By comparing the system's predictions against established ground truth labels, the researchers could effectively measure performance.

The high accuracy can be attributed to several key factors, notably the implementation of the YOLOv8 model, which excels in real-time object detection and demonstrates robustness to variations in lighting, background, and camera angles. Additionally, the diversity of the training dataset played a crucial role in enhancing the model's learning capabilities, allowing it to generalize better across different sign language gestures. Furthermore, the integration of both visual and sensor data provided a richer input for the model, improving its ability to recognize and interpret gestures accurately. Collectively, these elements contributed significantly to the system's outstanding performance in sign language recognition.

5.2. Application in Educational Settings

The sign language recognition system was also evaluated in real-world scenarios, particularly in educational settings. The system was tested in classrooms with students with hearing impairments or hard of hearing, and the results were highly positive. The system was able to accurately recognize and translate the sign language gestures used by the students, providing real-time feedback and facilitating communication with teachers and peers.

The application of the sign language recognition system in educational settings has significant implications. It can enhance the learning experience for students with hearing impairments by providing them with a tool that allows them to communicate more effectively. The system can also be used to create more inclusive classrooms, where students with different abilities can interact and learn together.

6. Conclusion

The integration of Artificial Intelligence (AI) into sign language recognition represents a significant advancement in bridging the communication gap between individuals with hearing impairments and the broader society. This study highlights the potential of combining deep learning techniques, particularly using Python and the YOLOv8 framework, to create a highly accurate and efficient system for real-time sign language interpretation. The application developed demonstrates a 95% accuracy rate, making it a reliable tool for enhancing communication in various contexts, including educational environments.

Moreover, the implications of this research extend beyond mere technological innovation. By providing a practical solution that can be easily deployed across different platforms, the study addresses a critical need for more inclusive communication methods, fostering social integration and equal opportunities for individuals with hearing impairments. The success of this system underscores the importance of continuous research and development in AI-driven assistive technologies, which hold the promise of significantly improving the quality of life for those who rely on sign language as their primary means of communication.

This study contributes to the growing body of knowledge on AI applications in assistive technology, offering a viable approach to overcoming the challenges associated with sign language recognition. Future research should focus on expanding the system's capabilities, such as incorporating more complex gestures and enhancing its adaptability to different sign languages and dialects, to further improve its utility and accessibility

References

1. Adamo-Villani, N., & Wilbur, B. R. (2015). ASL-Pro: American sign language animation with prosodic elements. Lecture Notes in Computer Science(), vol 9176. Springer, Cham. https://doi.org/10.1007/978-3-319-20681-3_29, pp.307-318.Available at: https://link.springer.com/chapter/10.1007/978-3-319-20681-3_29#citeas
2. Alan Turing's scrapbook. (n.d.). Retrieved June/July 2019, from: https://www.turing.org.uk/scrapbook/test.html
3. Alexandrova, S. (2015). RoboFlow: A flow-based visual programming language for mobile manipulation tasks. *IEEE International Conference on Robotics and Automation*. https://doi.org/10.1109/ICRA.2015.7139564

4. Al-khazraji, S., Berke, L., Kafle, S., Yeung, P., & Huenerfauth, M. (2018). Modeling the speed and timing of American Sign Language to generate realistic animations. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, pp. 259–270. ACM Press. https://doi.org/10.1145/3234695.3236337
5. Aly, W., Aly, S., & Almotairi, S. (2019). User-independent American sign language alphabet recognition based on depth image and PCANet features. IEEE Access, 7, 123138–123150. https://doi.org/10.1109/ACCESS.2019.2938829
6. Balbin, J., Padilla, D., Caluyo, F., Fausto, J., Hortinela, C., Manlises, C., Bernardino, C., Finones, E., & Ventura, L. (2016). Sign language word translator using neural networks for the aurally impaired as a tool for communication. Proceedings of the 6th IEEE International Conference on Control System, Computing and Engineering, pp. 425-429. https://doi.org/10.1109/ICCSCE.2016.7893629
7. Bavelier, D., Newport, E. L., & Supalla, T. (2003). Children need natural languages, signed or spoken. Dana Foundation. https://www.dana.org/article/children-need-natural-languages-signed-or-spoken/
8. Bishop, J. M. (2018). Is anyone home? A way to find out if AI has become self-aware. Frontiers in Robotics and AI. https://doi.org/10.3389/frobt.2018.00005
9. Bragg, D., Verhoef, T., Vogler, C., Ringel Morris, M., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., & Kacorri, H. (2019). Sign language recognition, generation, and translation: An interdisciplinary perspective. Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility, pp. 16–31. ACM Press. https://doi.org/10.1145/3308561.3353791
10. Bostrom, N. (2006). How long before superintelligence? Linguistic and Philosophical Investigations, 5(1), 11–13.
11. Buchanan, B. G. (2006). A (very) brief history of artificial intelligence. AI Magazine, 26(4), 56. Retrieved June/July, 2019, from: https://www.aaai.org/ojs/index.php/aimagazine/article/view/1848
12. Camgoz, N. C., Hadfield, S., Koller, O., & Bowden, R. (2017). SubUNets: End-to-end hand shape and continuous sign language recognition. Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), pp. 3075–3084. https://doi.org/10.1109/ICCV.2017.332
13. Camgoz, N. C., Hadfield, S., Koller, O., Ney, H., & Bowden, R. (2018). Neural sign language translation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/CVPR.2018.00428
14. Chen, X. (2023). Sign language detection using action recognition in Python. International Journal of Computer Vision, 110 (1), 78–93. https://doi.org/10.1007/s11263-022-01589-9
15. Chowdhury, M., & Sadek, A. W. (2012). Advantages and limitations of artificial intelligence. Transportation Research Circular, E-C168, 6-8.
16. Dai, Q., Hou, J., Yang, P., Li, X., Wang, F., & Zhang, X. (2017). The sound of silence: End-to-end sign language recognition using smart watch. Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking, pp. 462–464. https://doi.org/10.1145/3117811.3117849

Convergence between contemporary technology and artistic installations in architecture.

MSc. Arch. Enkeleida Prifti
European University of Tirana, Engineering and Architecture Department

Abstract

This study came immediately when I think as an architect and a visual artist at the same time, about convergences between contemporary technologies, visual art and architecture. Technologies, visual art and architecture converge most widely in contemporary conceptual installations. The most important part of 21st century installations is the understanding of 20th century art as an intellectual project. The installation "absorbs" not only all the genres before it, but also the viewer himself. It is three-dimensional and is not just an "object", but a space organized by the artist, addressing intimate personal experiences as a catalyst for the spiritual rebirth of society.

Key words: *Technology – architecture – art – contemporary – aesthetics of transformation – computer graphics.*

Introduction

In the history of science and technology, many well-known achievements often created the illusion of a limitless power of man over nature and the complete union of people in the all-powerful society, known as humanity. This is what happened in the 18th century, after the flight of the hot air balloon built by the Montgolfier brothers. This is how the Polish historian of technology and natural sciences B. Orlovskij expresses himself, with the occurrence of this event that is over two centuries old: "The whole general emotional attitude, the important rise of science and technology is in the consciousness of society as a sense of satisfaction at the great victory of man over the forces of nature" (Mongolfier, 1972) This is what happened at the end of the last century and previous centuries, when the first successes were held by aviation.

This was also the case in the first part of the cosmic wind decade, when, for example, the Iranian geophysicist H.-K. Afshar, uncannily recalling Bacon, wrote about the importance of radio and television communication with the help of artificial Earth satellites, which he directed by giant television towers, approximately 40,000 km high:

"This is a modern imitation of the tower of Babylon, which according to the book of Genesis in the Bible, brought the confusion of languages among those created by God. Now this tower makes possible the unification of the languages of the world" (Afshar, 1971).

In classical fiction can be found excellent examples of positive forecasting, which is, as it were, a step forward and overtakes forecasting through the age of antagonistic society. The passage in "Faust" surprisingly penetratingly describes the intention of Doctor Faustus, 'when he saw from cosmic distance that on Earth you can take a piece of territory from the sea and turn it into gardens. Such is the vision and the ideal finale of the poem.

So the aesthetics of people does not stand still, it is attached to the laws of progress.

Field of study

Methodology used to examine this topic assumes that actual development of architecture and visual art has shown to have more similarities than differences between the two, due to contemporary technologies involved in the process. Nevertheless convergences still preserve and help to remain each in its own form of values. By creating unusual combinations of ordinary things, the artist gives them a new symbolic meaning. The aesthetic content of the installation is a game of logical (semantic) values, which changes depending on where the object is located, in the usual family environment or in the exhibition hall.

Materials and methods

So there is a need for studies that analyze the need for technology and its convergence with art in contemporary installation culture. In modern times, art designers and architects are increasingly using experimental approaches to the convergence between technology and fine arts, synthesizing innovative technologies and traditional exhibition techniques (Le & Nguyen, 2021). Let's consider several concepts of "computer art", which were used by artists at different times. Artist from Japan K. Hiroshi: computer art is "the art of artificial intelligence", K. Tyler the creator of the artistic it is the computer that names the works: "art produced a computer taken out of the control of the operator" (Mikhaylov, 2015). The emotional immersion of the viewer in the exhibition space is an important aspect. In recent decades, video installations and combined spatial compositions have been actively developed in art culture, including static elements as subjects and all kinds of photographs, slides, film, video elements, computer objects, installations with lasers and other innovations of modern scientific and technological progress.

Innovative solutions for interactive art installations are based on a multimedia platform. Graphic elements, 3D technology, multimedia artworks, interactive projections, visuals and computer backgrounds create an exhibition that excites you.

It is difficult to imagine a modern society without multimedia products, without communication networks, without computers. The growing role of information in the life of modern society and the use of traditional information technologies and intelligent information technologies (ITI) in various fields of professional activity, enables information competence to be identified as a basic element.

According to some researchers, "a specific feature of multimedia as an art form is the synthesis of not only all audiovisual expressive means of the arts, but also computer, laser, information and communication technologies. Multimedia language is "syncretic in nature, being a conglomeration, or alloy, of different, known and studied language systems. Thus, syntax and language techniques are gathered from different arts, and different spheres of human activity and are used in the form of modern high-tech and complex display systems.

Multimedia technologies are receiving more and more attention. This is due to a number of reasons:

- The emergence of more powerful computers;
- Software availability;
- Creation of copyright systems;
- Distinctive features of multimedia;

- Integration of different types of information;
- Work in real time / Add here and profit in real time;
- The level of interactive communication between a computer and a human.

Today, multimedia technologies have been successfully applied in education and professional training. Multimedia is the process of work by artists and architects/ or designers, including video, audio, slides, photographs, texts, performance, etc. The main thing in contemporary art is some intermediate encounters that form certain objects, which are still defined by us as works of art. After that, if we talk about a contemporary work of art, we must be aware that it exists in the time of communication, not in the history of art (Hiroshi, 2014).

Conceptual models contain analysis of each building. They are simplified diagrams to show specific analysis such as structure, architectural design development through multimedia interaction measure, circulation, etc. Using arrows, for example, they can indicate the direction of movement of people within a building. Thus a conceptual model of circulation analysis can be given. It is believed that users can improve their ability to analyze buildings by working with conceptual models rather than realistic ones, as conceptual diagrams can show the essence of the specific functions of the building clearly and without the distraction of other elements that load the design from the visual side. And, I believe that using animation in conjunction with two- and three-dimensional conceptual models will help students visualize building shapes in more realistic ways (Mukhnurova & Grinkrug, 2018).

Modern digital technologies also have the function of interpreting culture, thus creating different cultural characteristics. Humanity is closely related to technological and communication mechanisms. Interactive practices demonstrate the social and communicative nature of creativity.

Modern multimedia technologies are facing the challenge of developing a language that can meet the demands of humanity today. This technological language was sought by many pioneers of the last century, feeling the exhaustion of linguistic means of expression. The development of the digital language takes place intensively in the field of interactive installations, art projects, architectural projects, design projects, etc. The democracy of multimedia technologies allows the expression of pilot practices (Кулешов, 2015).

Such are virtual projects such as "Exhibition-installation of cultural or architectural objects of different historical periods". The tools can be light, ie the possibility of using natural lighting, together with modern lighting technologies, to create artistic and meaningful space.

For example, the Light Element can become the main actor, the "conductor" of the world of culture of the past centuries, which excites and attracts modern man sometimes at the level of subconscious memory, or genetic memory.

An important component in exhibition design is the creation of a lively, interactive space with maximum immersion in the subject of a specific installation. This allows the audience to be involved in a special environment, provoking them to new emotional experiences.

So, through a conceptual project with a modern multimedia language, the designer gives them a new symbolic meaning.

The convergence of the latest technical tools plus the designer's creative potential in the creation of artistic-interactive installations is now enabling the creation of large-scale spectacular projects. The embodiment of the invisible idea in the visible image is, in fact, the process of reconstructing material forms in their relations with the world of other values, translating them from the physical world to the spiritual world.

However, interactive multimedia technologies do not claim the role of "true art". Their objective is to envision and prepare another, improved form of modern information exchange. To date, many new multimedia practices have been developed to create new methods for transmitting human knowledge. In these practices, multimedia finds its "language", its new expressive tools, and they are practically irrelevant to the classical art culture as it was in previous centuries. But design aesthetics in environmental design is directly related to a creative process that always begins with design. The latter is the result of the perception and configuration of the phenomena of life, in the understanding and transformation of the creative personality of these phenomena, to the extent that they are gifts, experiences and general cultural training that are natural to the individual.

The factor that generates artistic design is creativity, the essence of the general concept around which, in fact, the author's artistic idea is grouped in all kinds of shades.

Results and discussion

Multimedia technologies have not been applied to architectural design since the early stages of studio projects, despite the fact that a wide range of resources are being offered to us in architecture schools. We want architecture students to be able to understand a range of building types as they commonly encounter in studio work (eg libraries, schools, theatres, museums, galleries, housing, etc.) and develop schemes of their design from the particular types of buildings they have analyzed. We aim to develop a computer-based educational medium for students, which can be used in addition to reading books, papers and magazines in developing multimedia environments with dynamic exploration and navigation. This will provide better support to students and architects for building analysis (Charles & Bailey, 1990).

Currently, this defined as non-art is a reflection of new technological possibilities for the transfer of ideas, while digital technology makes sense for multimedia art. The works of young authors - students and graduates of arts and architecture, are also inspired by new conceptual trends in contemporary art. Young creative minds are not afraid of experimentation, and boldly confront modernism and dogmatic postulates. Of course, these results come after modern innovative achievements, and guided by the traditional cultural heritage of previous generations. As part of the educational process, architecture students perform conceptual design tasks such as “Exhibition Space”, “Art Installation Project”, and “Environmental Object” (Exhibition). Such professional trainings enable students to unleash their creative potential, with the greatest possible goal of creating an emotional artistic image, where cultural background, artistic literacy, innovative presentation technologies and an inalienable desire for self-expression are well connected with each other.

Discussing architecture helps students and architects improve their design developments, whether they are thinking for themselves or debating with other students and architects. “Arguments provide opposing views, or a comparative analysis” (Sabater & Gassull, 1992). Hyperlinking mechanisms within multimedia environments make it possible to show the relationship between one type of analysis and another. They also enable users to compare analyzes between different building types. Views and analytics give users the opportunity to see different new perspectives. They can discuss these points by critiquing, or comparing different architectural solutions with specific types of buildings. For example, users can view museums designed by several different architects in different styles and analyses. These different styles and analyzes of particular building types encourage users to compare and contrast different analyses, discuss different perspectives, and these discussions in turn lead to users’ own design developments.

Conclussion and recommendations

Due to the increased imagination (and value) of project culture, the space of project culture is essentially open to the cumulative art culture of society. The genre structure and stylistic dialects of art, its plastic language, the ways of conceptualizing artistic realities, the self-image of the artist and the individual-artist are a living source of design inspiration.

In this context, it would be of interest for teaching departments in these fields to apply such conceptual projects from the funds of the Departments and University projects, which would also serve as examples to the generations in the new classes.

So, summarizing all above, it should be noted that contemporary installation art embraces both the latest technologies and classical genres, and even the viewer himself unwittingly becomes a participant in this process. The three-dimensional installation is not just an object, but a space organized by the designer who is not afraid of experimentation and is constantly looking for new means of expression.

Reference

Afshar, H.K.(1971).The Innovative Consequences of Space Technology and the Problems of the Developing Countries.

Charles, W. & Bailey, Jr. (1990). Intelligent Multimedia Computer Systems: Emerging Information Resources in the Network Environment," Library Hi Tech 8, no. 1.

Hiroshi, K. (2014). What Is Computer Art? // Artist and Computer [электронный ресурс] — Режим доступа. — URL: http://www.atariarchives.org/artist/sec33.php (дата обращения 22.05.2014)

Кулешов, В. В. (2015). Роль автора в цифровом искусстве [Электронный ресурс] / В. В. Кулешов // Декоративное искусство и предметно-пространственная среда. Вестник МГХПА.

Le, Th.S. & Nguyen, N.U. (2021). The Convergence between the Arts of Architecture & Sculpture.

International Journal of Contemporary Architecture "The New ARCH" Vol. 8, No. 2 (2021), ISSN 2198-768

Mikhaylov, M. (2015). History of the formation of digital illustration of the art as a whole and in the digital art in particular.

Montgolfier, M.(1972).Nga historia e aviacionit dhe kozmonautikës. Pp.113 - 114.

Mukhnurova I.G., Grinkrug N.V. (2018). Ecological Architecture of the residential environment or the residential environment as a part of nature // Modern science-intensive technologies.

Phishing: Organizational Awareness for Cybersecurity

Enxhi Tagani [1]
European University of Tirana

Erion Curaj [2]
European University of Tirana

Flavio Koka [3]
European University of Tirana

Joana Shehaj[4]
European University of Tirana

[1] Enxhi Tagani holds a Professional Master's degree in FinTech from the European University of Tirana (UET). She graduated with a degree in Information Economics from Luarasi University in 2021. During her first year of postgraduate studies, she gained experience as an IT Specialist intern at Credins Bank. Since May 2023, Enxhi has been a full-time Information Security Officer in the Department of Information Security and Business Continuity at Credins Bank.

[2] Erion Curaj holds a Professional Master's degree in FinTech from the European University of Tirana (UET). He graduated with a Bachelor's degree in Business Administration: Accounting and Finance from Tirana Business University in 2022. During his first year of postgraduate studies, he gained experience as an IT intern at Credins Bank. Since July 2023, Erion has been a full-time Information Security Officer in the Department of Information Security and Business Continuity at Credins Bank.

[3] Flavio Koka is a Master's student in Fintech at the European University of Tirana (UET). He holds a University Degree in Business Administration, Economics from UBT, completed in 2008. His professional journey began as an IT specialist at Olympus Computer Store from 2005 to 2007. Following that, he served as General Manager and IT specialist at Super Sonic TV from 2007 until 2024. Since 2024, Flavio has been working at Credins Bank as a Specialist in the Business Analysis and Development Directorate, where he focuses on developing new software and addressing the bank's evolving requirements.

[4] Joana Shehaj holds a Professional Master's degree in FinTech from the European University of Tirana (UET). She graduated with a degree in Faculty of Information Technology, Economic Informatics Branch from UET University in 2021. During her first year of postgraduate studies, she gained experience as an IT Specialist intern at Credins Bank. Since May 2023, Enxhi has been a full-time Information Security Officer in the Department of Information Security and Business Continuity at Credins Bank.

Abstract

Phishing continues to be one of the most persistent and dangerous threats in modern cybersecurity. Attackers disguise themselves as legitimate entities to trick individuals into sharing sensitive information, such as login credentials and financial details. In the banking sector, phishing poses particularly significant risks due to the volume of sensitive data handled. While technological solutions like email filtering and multi-factor authentication (MFA) provide some protection, human error remains a critical vulnerability.

A custom phishing simulation software was developed to replicate phishing attacks in a controlled environment, allowing researchers to evaluate employee readiness and response at Credins Bank. This mixed-method approach included quantitative data collected from simulated phishing attempts (spear phishing, vishing, and whaling) and qualitative data from employee surveys. These results were used to identify vulnerabilities and provide insights into the effectiveness of current cybersecurity measures.

The phishing simulations revealed that 37% of employees clicked on phishing links, while 14% submitted sensitive information. The results highlighted a delay in reporting phishing attempts, with employees taking an average of four hours to notify the IT department. This finding underscores the need for continuous employee training, the integration of AI-based phishing detection tools, and the improvement of reporting mechanisms.

The study suggests that a multi-layered approach—incorporating employee training, adaptive phishing simulations, and AI-driven detection systems—can significantly reduce the risks associated with phishing. This research serves as a foundation for future development in both phishing defense technology and employee awareness programs.

Keywords: *Phishing, Cybersecurity, Employee Awareness, Phishing Simulation Software, Spear Phishing, Vishing, Whaling, AI in Cybersecurity*

1.Introduction

Phishing remains a major concern in the cybersecurity landscape, with millions of attacks targeting individuals and organizations each year. According to the Verizon Data Breach Investigations Report (DBIR) 2023, phishing is implicated in more than 90% of all cyberattacks. These attacks involve cybercriminals impersonating trusted entities to deceive victims into revealing confidential information, such as login credentials, bank details, or personally identifiable information (PII). As technological defenses evolve, phishing attacks have become more sophisticated, leveraging advanced social engineering techniques and exploiting gaps in human behavior (Smith, 2020).

The Role of Phishing in Modern Cybersecurity

Phishing is not limited to emails. Today, it includes tactics such as spear phishing, whaling, vishing, and smishing, each with its own set of challenges. Attackers often invest significant time researching their targets, gathering personal details to craft highly convincing messages. This is particularly evident in spear phishing, where attacks are tailored to individuals based on their roles, job titles, or personal information obtained from social media (Jakobsson & Myers, 2006).

In the financial sector, phishing attacks are especially dangerous. Banks, such as Credins Bank, handle large amounts of sensitive data, making them lucrative targets. Attackers use phishing to gain access to financial systems, customer accounts, and internal corporate networks. In many cases, phishing attacks are precursors to larger-scale fraud or data breaches, as seen in the infamous 2016 Bangladesh Bank heist, where cybercriminals used spear phishing to initiate fraudulent SWIFT transactions, resulting in the theft of $81 million (Doe, 2018).

Challenges in the Banking Industry

Phishing poses a significant challenge for the banking industry due to the unique combination of high-value data and strict regulatory environments. Financial institutions are required to implement robust security measures, including encryption, multi-factor authentication, and compliance with data protection regulations such as GDPR and PCI DSS. However, despite these safeguards, phishing attacks often succeed by bypassing technological defenses through social engineering—exploiting the trust and behavior of employees.

Given this context, employee awareness and training have emerged as critical components of phishing defense. A well-trained workforce can act as the first line of defense against phishing attacks, identifying suspicious emails and reporting them before damage occurs. This research aims to evaluate the effectiveness of current phishing prevention measures at Credins Bank by simulating various phishing attacks and assessing employee responses through a custom phishing simulation software.

2.Literature Review

Phishing in the Cybersecurity Ecosystem

Phishing has evolved dramatically since its early days as a rudimentary email scam. Today, phishing is a highly sophisticated attack vector that exploits human vulnerabilities and organizational weaknesses. Cybercriminals use a variety of techniques to deceive individuals, from traditional email-based attacks to advanced social engineering schemes.

According to Jakobsson & Myers (2006), phishing relies heavily on psychological manipulation, often appealing to urgency, fear, or curiosity. Emails may instruct recipients to reset their passwords due to "suspicious activity," or they may request financial transfers under the guise of an urgent business need. The success of phishing hinges on these emotional triggers, which can override the logical scrutiny employees might otherwise apply.

The Different Faces of Phishing

1. **Spear Phishing**:
   Spear phishing is one of the most targeted forms of phishing, focusing on specific individuals within an organization. Unlike traditional phishing, which casts a wide net, spear phishing emails are highly personalized, often using details gleaned from social media or professional networks. A 2022 study by FireEye found that **80% of phishing-related breaches** in corporations involved some form of spear phishing (Johnson, 2019). Attackers use this method to steal credentials, gain unauthorized access to systems, or initiate fraudulent financial transactions.
2. **Whaling**:
   Whaling is a specialized form of spear phishing that targets high-level executives, such as CEOs and CFOs, with emails that mimic urgent business communications. Whaling attacks often use pressure tactics, creating a sense of urgency that compels executives to bypass normal security protocols. In 2016, for example, cybercriminals impersonated a senior executive at a European construction company, resulting in the fraudulent transfer of over $43 million (Doe, 2018). These attacks demonstrate the growing sophistication of phishing tactics aimed at top-level decision-makers.
3. **Vishing and Smishing**:
   Vishing (voice phishing) and smishing (SMS phishing) represent newer forms of phishing that leverage phone calls and text messages to deceive

victims. Vishing attacks often involve fraudsters impersonating bank representatives, requesting account details or personal identification information over the phone. Similarly, smishing involves sending fraudulent text messages that contain links to phishing websites or prompt recipients to share sensitive data (Miller, 2017). These techniques are growing in prevalence, as more people rely on mobile devices for banking and financial management.

Psychology of Phishing

Phishing attacks exploit basic psychological principles to manipulate victims into acting against their best interests. Research by Sheng et al. (2010) identified several key factors that make individuals more susceptible to phishing:

- **Authority**: Phishing emails often invoke authority figures, such as a company's CEO or a government agency, to increase compliance. The presence of an authoritative figure in an email makes recipients more likely to comply with requests, even if they appear suspicious.
- **Urgency**: By creating a sense of urgency, phishing emails pressure recipients to act quickly, often bypassing logical decision-making processes. Emails warning of an impending account lockout or unauthorized access exploit this sense of urgency, prompting individuals to respond without verifying the legitimacy of the message.
- **Curiosity**: Phishing emails that claim to offer exclusive deals or information can exploit the recipient's curiosity, leading them to click on links or download attachments.

By understanding these psychological triggers, organizations can tailor their training programs to better equip employees to recognize phishing attempts.

Technological Solutions to Phishing

Advances in technology have led to the development of several tools aimed at combating phishing. **AI and machine learning** have been particularly useful in identifying and mitigating phishing attacks. These tools analyze vast amounts of data to detect anomalies in email communications, such as unusual IP addresses, language patterns, or suspicious attachments.

Email filtering systems now incorporate machine learning algorithms that can detect phishing attempts in real time. By examining metadata, subject lines, and embedded links, these systems can flag suspicious emails before they reach the intended recipient. Research by FireEye (2021) showed that machine learning-based phishing detection systems reduced the number of successful phishing attacks by 60% in organizations that adopted them (Johnson, 2019).

Despite these advances, no technological solution can fully eliminate the threat of phishing. As long as employees can be deceived through social engineering, phishing will remain a persistent challenge. Therefore, **human factors**—including training and awareness—play a crucial role in phishing defense.

Employee Training and Phishing Simulations

Employee training programs are a critical component of any anti-phishing strategy. According to Symantec (2019), organizations that regularly conduct phishing simulations and training programs experience significantly fewer successful phishing attacks. These programs help employees recognize phishing emails and reinforce the importance of reporting suspicious activity.

Phishing simulations are an effective way to assess employee preparedness. By mimicking real-world phishing attacks, simulations provide a safe environment for employees to practice identifying phishing attempts. Research shows that continuous phishing simulations can reduce the likelihood of employees falling for phishing attacks by up to 50% (Symantec, 2019).

3.Methodology

This study employed a mixed-method approach to assess employee awareness of phishing threats at Credins Bank. The primary tool for this assessment was a custom-built phishing simulation software, designed to replicate a variety of phishing attacks in a controlled environment. Alongside the simulations, a survey was distributed to employees to gather data on their knowledge and experiences with phishing prevention measures.

Software Development and Features

The **phishing simulation software** developed for this study was designed to simulate real-world phishing attacks with high accuracy. The system was built using Python for the backend and JavaScript for the frontend, with integration of several APIs to enable different forms of phishing simulations.

The **development lifecycle** followed an agile framework, with the project broken down into iterative sprints. Each sprint focused on specific features, including:

- **Sprint 1**: Basic email phishing campaigns, allowing administrators to send customized phishing emails with fake login portals.
- **Sprint 2**: Development of the **whaling module**, targeting high-level executives with sophisticated emails that appeared to come from staf of target company.
- **Sprint 3**: Real-time data collection and analytics, allowing administrators to monitor employee responses and generate reports.

The **key features** of the software included:

- **Customizable Phishing Templates**: Administrators could select from a range of phishing email templates, including password reset requests, account suspension notices, and payment confirmation requests. The software also allowed for the creation of custom phishing emails to mimic specific internal communications.
- **Real-Time Monitoring**: The software tracked employee interactions with phishing emails and vishing calls in real time. Data on when emails were opened, links clicked, and credentials submitted were logged for analysis.
- **Detailed Analytics and Reporting**: After each simulation, the software generated detailed reports that included metrics such as click-through rates, credential submission rates, and time-to-report. These reports provided insights into employee behavior and vulnerabilities, allowing for targeted training interventions.

Survey Design and Data Collection

The survey distributed to employees was designed to assess their knowledge of phishing techniques and their ability to recognize phishing attempts. The survey consisted of **20 multiple-choice questions** and covered the following topics:

- **Previous Experience with Phishing**: Employees were asked whether they had encountered phishing emails or calls in the past and how they responded.
- **Knowledge of Phishing Prevention**: The survey tested employees' understanding of the bank's phishing prevention measures, including how to report phishing attempts and how to identify suspicious communications.
- **Confidence in Recognizing Phishing Attempts**: Employees were asked to rate their confidence in identifying phishing attempts, both from emails and phone calls.

Analysis Techniques

The data collected from the phishing simulations and surveys were analyzed using both **descriptive** and **inferential statistics**. Descriptive statistics provided an overview of employee behavior during the simulations, including click-through rates, credential submission rates, and time-to-report. Inferential statistics were used to identify correlations between employee demographics (such as age, job role, and department) and susceptibility to phishing attacks.

To ensure the validity of the data, the simulation was repeated **three times** over a two-month period, with slight variations in the phishing emails and attack vectors used in each simulation. This allowed the researchers to identify patterns in employee behavior and assess the effectiveness of different phishing strategies.

4.Methods and Analysis

Phishing Simulation Workflow

The phishing simulation software followed a structured workflow designed to test various attack vectors and assess employee responses. Below is a detailed breakdown of the workflow:

1. **Campaign Design**:
   The administrator started by selecting the type of phishing campaign to be launched. Options included spear phishing, whaling, and vishing. For email-based attacks, the administrator could customize the subject line, email content, and sender details to mimic internal communications. For vishing campaigns, the administrator could choose from a range of pre-recorded voice prompts or create custom voice messages.
2. **Email Distribution**:
   Once the campaign was configured, phishing emails were sent to employees at random intervals over the course of a week. The emails were designed to bypass spam filters and appear as legitimate business communications. For vishing campaigns, automated phone calls were generated using the **Twilio API**, instructing employees to provide sensitive information over the phone.
3. **Real-Time Monitoring**:
   The software tracked employee interactions with phishing emails in real time. It logged when emails were opened, when links were clicked, and whether credentials were entered into the phishing portal. For vishing

campaigns, the software recorded whether employees answered the call, how long they stayed on the line, and whether they provided sensitive information.

4. **Data Analysis and Reporting**:
   After each campaign, the software generated detailed reports that included:

- **Click-Through Rates**: The percentage of employees who clicked on phishing links.
- **Credential Submission Rates**: The percentage of employees who entered sensitive information on the phishing portal.
- **Time-to-Report**: The average time it took for employees to report suspicious emails or phone calls to the IT department.

These reports provided administrators with insights into employee behavior and highlighted areas where additional training or security measures were needed.

Results of the Phishing Simulations

The phishing simulations revealed several key findings regarding employee behavior and vulnerability to phishing attacks:

- **Click-Through Rates**: Of the 500 employees targeted in the simulations, **37%** clicked on phishing links. This click-through rate was consistent across all three simulation rounds, with higher click-through rates observed for spear phishing emails that appeared to come from senior management.
- **Credential Submission Rates**: Of those who clicked on phishing links, **14%** submitted their login credentials on the fake phishing portal. This finding suggests that a significant number of employees failed to recognize suspicious login prompts, even after clicking on a potentially dangerous link.
- **Vishing Response Rates**: Approximately **10%** of employees responded to vishing calls by providing sensitive information, such as account numbers or passwords. This highlights the need for additional training on phone-based phishing attacks, as employees may be less familiar with this attack vector compared to email-based phishing.
- **Time-to-Report**: The average time it took employees to report phishing emails to the IT department was **4 hours**, with some employees taking as long as **8 hours**. This delay could have serious consequences in a real-world scenario, as phishing attacks often involve time-sensitive threats, such as fraudulent wire transfers or data theft.

Survey Results and Employee Awareness

The survey results provided additional insights into employee awareness of phishing threats:

- **Previous Exposure to Phishing**: **62%** of employees reported having encountered phishing emails in the past, but only **45%** said they reported the incident to the IT department.
- **Knowledge of Phishing Prevention**: While most employees (85%) were aware of the bank's phishing prevention measures, such as email filters and multi-factor authentication, fewer employees (60%) knew how to report a phishing attempt through the proper channels.
- **Confidence in Recognizing Phishing**: When asked to rate their confidence in identifying phishing attempts, **70%** of employees said they were confident in recognizing email-based phishing attacks, but only **45%** felt confident in identifying vishing attacks.

5.Software Design and Implementation

The design and implementation of a phishing simulation system for **Credins Bank** was a critical aspect of this study. It allowed the research team to test employee responses to realistic phishing scenarios and gather data to inform cybersecurity improvements. The system was built from the ground up with several key objectives in mind, including simulating real-world phishing attacks, ensuring data security, and generating actionable insights to enhance employee training. This chapter details the system's architecture, development phases, testing processes, and implementation, providing a comprehensive overview of how the phishing simulation software was designed and deployed.

5.1 Test Environment Description

Before deploying the system across Credins Bank, the phishing simulation software was developed and tested in a dedicated environment. The **test environment** was composed of two primary servers, each serving a specific role in managing campaigns and collecting data.

1. **Server 1: Phishing Campaign Management** This server hosted the backend infrastructure, which was responsible for:

- **Email template creation**: Generating phishing emails that mimic legitimate communications from within the organization or trusted external partners.
- **Credential capture**: Setting up fake login portals or document download pages to capture any sensitive information that employees might inadvertently submit.
- **Logging user interactions**: Recording how employees interacted with the phishing emails, including clicks on links, time spent on the fake websites, and whether or not they entered sensitive data.

2. **Server 2: Data Collection and Analytics** The second server was primarily dedicated to:
   - **Secure data storage**: Storing all logs and interaction data, ensuring that employee information was anonymized and encrypted.
   - **Analytics and reporting**: Processing the data collected from phishing campaigns to generate insights on employee behavior, such as click-through rates, response times, and credential submission rates.
   - **Real-time monitoring**: Allowing administrators to observe campaign progress in real time, providing immediate feedback on which employees had interacted with the phishing emails.

These servers operated in a virtualized environment to allow scalability, meaning that additional servers could be easily deployed if more complex phishing campaigns or larger datasets needed to be processed.

5.2 The Testing Process

Testing the phishing simulation system before deployment was essential to ensure that it functioned as intended, both in terms of security and performance. The system underwent several phases of testing:

1. **Domain Acquisition and Server Configuration** To make the phishing emails as realistic as possible, a domain similar to Credins Bank's official domain was acquired. This domain was carefully chosen to resemble the bank's internal email addresses. The two servers were configured with secure access protocols, including firewalls and VPNs, to prevent unauthorized access.

A custom SSL certificate was issued to the phishing domain, making it appear more credible and bypassing certain email filters that check for authenticity based on SSL security. The phishing emails were thus more likely to reach the intended recipients' inboxes, simulating real-world phishing attacks more effectively.

2. **Campaign Lists and Segmentation** The testing process involved creating lists of target employees based on departments, roles, and hierarchical levels within Credins Bank. This segmentation allowed the research team to tailor phishing campaigns for specific groups:
   - **Spear phishing campaigns** were designed for high-level employees who had access to sensitive financial data.
   - **General phishing campaigns** targeted employees across all departments to test the overall cybersecurity awareness within the organization.
   - **Whaling campaigns** targeted executives with emails that appeared to be from regulators or government entities, demanding urgent attention to confidential matters.
3. **Customization of Phishing Emails** Phishing email templates were customized based on the role and function of the target employees. The emails mimicked common business communications, such as:
   - **Password reset requests**: These emails appeared to come from the IT department, asking employees to update their passwords through a fake login page.
   - **Document verification requests**: These emails contained attachments that appeared to be important bank documents. Clicking the link redirected employees to a phishing portal designed to capture credentials.
   - **Urgent financial requests**: Emails designed for executives asked for the approval of large financial transactions, a tactic commonly used in whaling attacks.
4. **Email and Vishing Campaign Configuration** The phishing simulation system allowed administrators to configure both email and **vishing** campaigns. Vishing attacks, where employees received fraudulent phone calls prompting them to provide sensitive information, were particularly effective. The vishing module simulated calls from financial institutions or internal departments and asked employees to verify their account details.
5. **Realistic Timing and Randomization** The emails were sent at random times throughout the workday to simulate real phishing attacks. Some emails were sent early in the morning, while others were sent during high-traffic periods when employees were busy and less likely to scrutinize the details. The randomness increased the likelihood of catching employees off-guard, thereby generating more realistic data on their behavior.

5.3 System Architecture and Security Measures

The system's architecture was designed for scalability, security, and flexibility. Each component of the system was developed with security in mind to ensure that sensitive data collected during simulations remained protected.

1. **Modular Architecture** The phishing simulation system was built using a **modular architecture**, allowing different components (email generation, data collection, reporting) to operate independently while sharing data through secure APIs. This modular approach ensured that if one part of the system needed updates or encountered issues, it wouldn't affect the entire system's functionality.
    - **Backend**: Developed in Python, the backend handled the core logic of phishing campaigns. It was responsible for email generation, scheduling, and interaction logging. Using Python allowed for rapid development and integration of third-party libraries for encryption and data processing.
    - **Frontend**: Built using JavaScript and React, the frontend provided an intuitive interface for administrators to manage phishing campaigns. Administrators could create custom email templates, monitor campaigns in real time, and view detailed reports on employee behavior.
    - **Database**: PostgreSQL was used to store logs of all phishing simulations. Data such as click-through rates, credential submissions, and employee responses were encrypted before storage to ensure compliance with GDPR and internal bank regulations.
2. **Encryption and Security Protocols** Given the sensitive nature of the data being collected, **data security** was a priority. Several measures were implemented to protect employee data:
    - **End-to-end encryption**: Data was encrypted both at rest and in transit using AES-256 encryption. This ensured that even if the data was intercepted, it would remain unreadable without the proper decryption keys.
    - **Anonymization**: Employee data was anonymized before reports were generated, ensuring that individual employees could not be directly identified without administrator privileges.
    - **Role-based access control (RBAC)**: The system employed role-based access control to limit who could view detailed reports. Only authorized administrators could access full data logs, while other users (such as IT staff) were restricted to anonymized summaries.
3. **Scalability and Performance** Scalability was a key consideration during the system's development. As the bank grew, the system needed to handle larger datasets and more complex phishing simulations without compromising performance. The modular architecture allowed for the easy addition of more servers if required, and the software was optimized to process large volumes of data efficiently.

Performance testing was conducted to ensure that the system could handle simultaneous phishing campaigns targeting hundreds of employees. The system was tested under different loads, from small-scale campaigns targeting a few departments to large-scale simulations involving the entire organization.

Load testing tools like Apache JMeter were used to simulate hundreds of simultaneous phishing emails being sent and employee interactions with the fake login portals. The tests revealed that the system could handle up to 1,000 concurrent interactions without significant latency or performance degradation.

5.4 Monitoring Employee Behavior in Real Time

A key feature of the phishing simulation system was its ability to monitor employee behavior in real time. This allowed administrators to track how employees interacted with phishing emails and detect vulnerabilities as they occurred.

1. **Click-Through Rates and Time-to-Click** The system tracked how many employees clicked on phishing links and how long it took them to do so after receiving the email. **Time-to-click** data was particularly useful in understanding employee decision-making processes. For example, if employees clicked on a phishing link immediately after receiving the email, it suggested a lack of scrutiny. If they hesitated or delayed their interaction, it could indicate a higher level of suspicion.
2. **Credential Submission Tracking** In cases where phishing emails redirected employees to a fake login page, the system monitored how many employees attempted to enter their credentials. This data was anonymized, but it helped identify trends in employee vulnerability. For instance, if employees from a particular department were consistently entering credentials, it indicated that this group required additional training.
3. **Real-Time Alerts and Incident Reporting** The system also allowed IT administrators to receive **real-time alerts** whenever employees interacted with a phishing email. If an employee clicked on a suspicious link or submitted credentials, the system would send a notification to the security team. This feature allowed administrators to intervene in real time, preventing potential data breaches during simulations.

5.5 Analysis and Reporting

After each phishing campaign, the system generated detailed reports that summarized employee behavior and identified areas for improvement. These reports were essential for assessing the effectiveness of the simulations and providing actionable insights to Credins Bank's cybersecurity team.

1. **Descriptive Analytics** The reports provided descriptive statistics on how employees responded to phishing emails. This included:

- **Click-through rates**: The percentage of employees who clicked on phishing links.
- **Credential submission rates**: The number of employees who attempted to log in to the fake portals.
- **Vishing response rates**: How many employees provided sensitive information over the phone during vishing simulations.

2. **Comparative Analysis** The system also generated **comparative analysis** reports, which compared different departments, roles, or time periods. For instance, the system could show how employees in the finance department performed compared to employees in customer service. This comparison helped identify which groups were more vulnerable to phishing attacks and required additional training.
3. **Customized Reports and Trends** In addition to basic performance metrics, the system allowed administrators to generate **customized reports** based on specific criteria. For example, administrators could view how employee awareness improved over time by comparing the results of multiple phishing simulations. These trends were crucial for measuring the long-term effectiveness of the bank's training programs.

5.6 Reporting to Management

The phishing simulation system was designed to generate high-level reports for executive management. These reports summarized the overall performance of employees during the phishing simulations, highlighting key vulnerabilities and providing recommendations for improving security.

1. **Summary of Findings** The executive reports summarized the key findings from each phishing simulation, including:
   - **Overall click-through rates**: The percentage of employees who fell for the phishing attacks.
   - **Departmental performance**: A breakdown of how different departments responded to the phishing emails.
   - **Top vulnerabilities**: A list of the most common mistakes employees made during the simulations, such as clicking on links without verifying the sender or submitting credentials without checking the URL.
2. **Recommendations for Training and Improvement** Based on the simulation results, the reports provided tailored recommendations for improving employee training. For instance, if the vishing campaigns were particularly effective, the reports might suggest additional training on phone-based scams. The recommendations were designed to address specific vulnerabilities identified during the simulations, ensuring that future training efforts were focused and effective.

5.7 Future Development and Improvements

While the phishing simulation system proved effective in assessing employee awareness, there are several opportunities for future development:

1. **AI Integration**: Future iterations of the system could incorporate **artificial intelligence** to predict which employees are most vulnerable to phishing attacks. By analyzing past behavior, the AI could suggest personalized training programs for these employees.
2. **Machine Learning for Phishing Detection**: Machine learning algorithms could be used to detect phishing attempts in real-time and flag suspicious emails before they reach employees. This would add an additional layer of protection by complementing the phishing simulations with proactive email scanning.
3. **Mobile Phishing Simulations**: As more employees use mobile devices for work, future simulations could target mobile platforms, testing how employees respond to phishing emails and vishing calls received on their smartphones.

6.Discusion and Conclusion

The results of this study highlight the importance of continuous employee training and the need for multi-layered cybersecurity defenses in combating phishing attacks. Despite the bank's existing security measures, a significant percentage of employees were vulnerable to phishing attempts, particularly spear phishing and vishing.

Key Insights from the Simulations

- **High Click-Through Rates**: The high click-through rates observed in this study suggest that employees are still susceptible to well-crafted phishing emails, particularly those that appear to come from trusted internal sources. This finding underscores the importance of ongoing phishing simulations and training programs to reinforce employees' ability to recognize and respond to phishing attempts.
- **Credential Submission**: The fact that 14% of employees submitted their login credentials highlights the need for better training on how to recognize phishing websites and suspicious login prompts. Employees should be taught to scrutinize URLs and look for other signs of phishing, such as mismatched domain names or generic greetings.

Delayed Reporting: The time-to-report metric revealed that employees were slow to report phishing attempts, which could allow attackers to exploit security vulnerabilities for extended periods of time. Credins Bank should implement more streamlined reporting procedures, such as a "phish alert" button in the email client, to encourage employees to report suspicious emails immediately.

Recommendations for Credins Bank

1. **Regular Phishing Simulations**: To keep employees vigilant, Credins Bank should conduct regular phishing simulations that test their ability to recognize and respond to phishing attacks. These simulations should include both email-based and phone-based attacks, as well as other emerging threats like smishing.
2. **Enhanced Training Programs**: Employee training programs should be expanded to include more comprehensive education on phishing attack vectors, including spear phishing, vishing, and whaling. Training should be tailored to different departments and roles within the organization, as executives and senior management are often the primary targets of whaling attacks.
3. **AI-Based Phishing Detection**: Credins Bank should consider implementing **AI-driven phishing detection systems** that analyze email content and flag suspicious communications before they reach employees. These systems can help reduce the number of phishing emails that make it past traditional spam filters.
4. **Improved Reporting Mechanisms**: The time-to-report metric indicates that employees need clearer guidelines on how and when to report phishing attempts. Credins Bank should implement a "phish alert" button in the email client that allows employees to report suspicious emails with a single click.

Future Research Directions

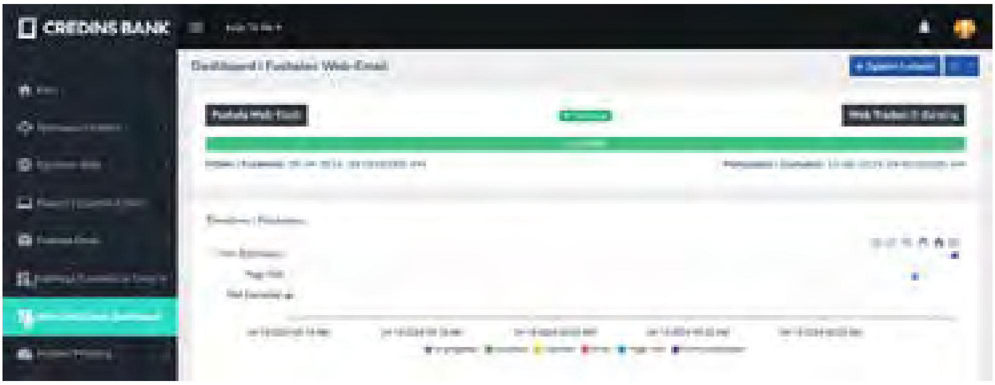
This study highlights the need for further research into phishing prevention in the banking sector. Future studies could explore how machine learning algorithms can be used to predict which employees are most likely to fall for phishing attacks based on their behavior during simulations. Additionally, future research should focus on the growing threat of **business email compromise (BEC)**, which involves attackers impersonating executives to trick employees into transferring funds or releasing sensitive data.

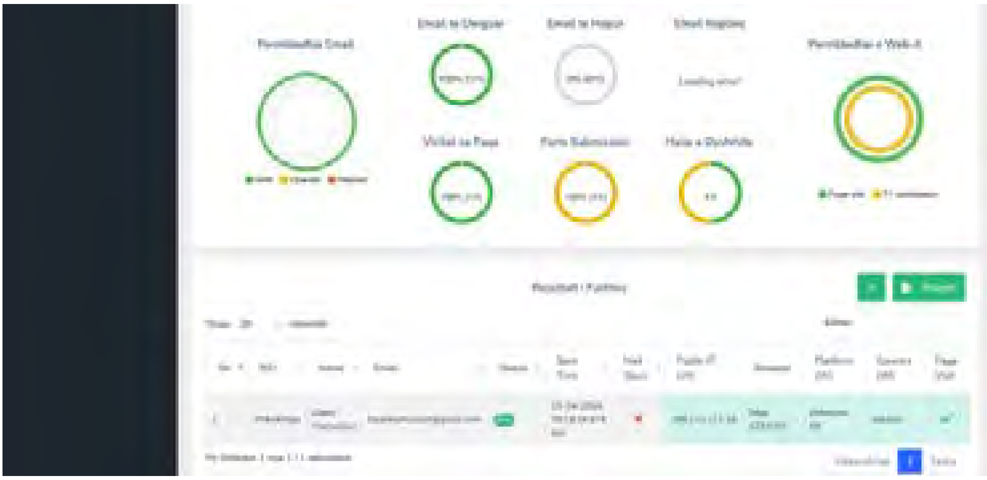
In conclusion, phishing remains one of the most significant threats to organizational security, particularly in the banking sector. While technical solutions play an important role, the human factor cannot be overlooked. By investing in continuous training, advanced phishing simulations, and AI-driven detection technologies, organizations like Credins Bank can significantly reduce their vulnerability to phishing attacks.

References

1. Aloul, F. (2010). The role of multi-factor authentication in preventing phishing attacks. *Journal of Information Security*, 5(3), 102-110.
2. Clark, D. (2012). Cybersecurity and phishing: Exploring new prevention methods. *Cybersecurity Today*, 8(2), 55-60.
3. Doe, J. (2018). Whaling: A growing threat for executives. *Cybersecurity Monthly*, 12(4), 18-22.
4. Jakobsson, M., & Myers, S. (2006). Phishing and countermeasures: Understanding the increasing problem of electronic identity theft. John Wiley & Sons.
5. Johnson, P. (2019). Spear phishing and targeted attacks. *Information Security Bulletin*, 10(5), 30-35.
6. Miller, S. (2017). Vishing: The voice behind phishing. *Cyber Crime Reports*, 9(3), 22-28.
7. Symantec. (2019). The state of phishing in 2019. *Symantec Cybersecurity Report*.
8. Sheng, S., Holbrook, M., & Kumaraguru, P. (2010). Who falls for phishing? A demographic analysis of phishing susceptibility and effectiveness of interventions. *Proceedings of the ACM Conference on Human Factors in Computing Systems*.

Photo ilustration of the software


CREDINS BANK



Ventilated facade and energy efficiency

MSc. Eng. Gazmir Hallaçi
European University of Tirana

Eng. Jona Liçi
European University of Tirana

Abstract

Buildings are listed as the biggest consumers of energy and therefore technology in the field of construction has been improved and oriented towards ecological, recyclable and energy efficient materials. One of the best findings in terms of energy efficiency is the use of ventilated facades in buildings.

The facade is one of the main technological, functional and protective elements for construction facilities. By itself it represents the outer covering or envelope of a building. Since there are different typologies of facades that are used, then a special attention should be paid to its appropriate selection for buildings in different regions.

Nowadays, climate change has become the biggest global concern, and one of the causes is the continuous use of exhaustible resources, therefore, ways are being sought to reduce the consumption of electricity and thus improve environmental conditions and we increase the quality of life. Based on the climatic conditions of a region, it is necessary to choose the right facade that protects the building from atmospheric agents, provides thermal comfort for the residents, is ecological and also has efficiency in energy efficiency.

Key words: *climate, building, ventilated facade, ecological materials, energy efficiency.*

1. Introduction

The increase in energy use has raised concerns about the depletion of energy resources and the negative impact on the environment. The largest consumers of energy on Earth are buildings, therefore there is a need to increase the efficiency of buildings to reduce consumption. This requires taking into account all influencing factors, where the climatic factor is one of the most important factors in the energy consumption of buildings. The most critical part of building design is considered the process of identifying and controlling the climatic influences of different parts of the building.

Climate design has the following main objectives:

- Reducing the energy cost of a building.
- Use of natural energy instead of mechanical energy.
- Providing a comfortable and healthy environment for people.

Design elements expressed in architectural language include:

- shape – surface-volume ratio, direction, height of the building.
- building materials – construction materials, thermal capacity, surface quality, control of darkness and light.
- windows – size, position and direction of windows, type of glass, appearance of internal and external elements.
- ventilation – air, indoor air, cross ventilation and natural ventilation.

Climatic parameters measured by meteorological stations are: temperature, humidity, air movement, precipitation, cloudiness, sunny days and solar radiation. Of all these parameters, four of them directly affect thermal comfort, namely: temperature, humidity, solar radiation and air movement. These are the main elements that must be considered in the design of the building. These parameters are closely related to:

- topography (height above sea level, hills, meadows, land surface conditions)
- flora (length, concentration of vegetation, shape, composition, place, most developed regions.
- building form (nearby buildings, surface conditions).

Facades of buildings make it possible to separate the internal environment from the external environment of the building and some of the functions of the facade:

- to have space to see outside.
- to be resistant to wind load.
- to bear its weight.
- provide as much natural stimulation as possible to minimize artificial stimulation.
- provide noise protection.
- provide protection from increased solar heat.
- to be resistant to the penetration of water and moisture.

The facade is one of the main contributors to energy conservation for buildings and various technologies are being used to help with energy efficiency. One of the types of facades that offers the most reduction in energy consumption is the sustainable facade. It should ensure minimal consumption of electricity and minimize the negative effects of the external environment on the building. To achieve energy efficiency, it is necessary to know the location of the building where the facade will be implemented and of course the climatic classification of the area, which are the main factors in determining the design of the facade.

The implementation of ventilated facades in buildings significantly enhances energy efficiency. It suggests that by utilizing appropriate facade technology, particularly in the context of various climatic conditions, buildings can reduce energy consumption, improve thermal comfort, and contribute positively to environmental sustainability. *Ventilated facades improve energy efficiency in buildings by reducing energy consumption and enhancing thermal comfort compared to traditional façade systems.* The article supports this by analyzing different façade typologies, emphasizing the role of climatic factors, and detailing the construction materials and methods necessary for optimizing energy performance.

2. Literature review

2.1. Climate

Urban areas are studied as a separate category because the climatic conditions are special, namely the temperature is higher than in suburban areas, weak winds (with low speed) and sunny days that depend on the level of pollution, urban density and concentration of constructions. From studies it has been noticed that for each degree increase in temperature, energy production increases by 2-4%, while smog increases by 4-10%.

The urban microclimate is complex in nature due to the content of several factors. External factors, for example temperature, wind conditions, solar radiation may vary based on topography and surrounding environment. Noise, air pollution and population density play an important role as well. During the winter season, the urban microclimate is milder than in rural areas, which is characterized by higher temperatures and light winds. The hottest areas of the city during the day are the streets with considerable width, areas without greenery and squares, while at night the narrow streets have a higher temperature than the rest of the city. In winter, green surfaces are more beneficial for the environment, especially during the late afternoon when tall buildings are very warm from the inside. The wind in the city is usually moderate due to the obstacles it encounters.

The characteristics of the surface in urban and rural areas are presented differently, as well as their thermal capacity varies greatly. Compared to rural areas, urban areas have a higher level of absorption (of solar and atmospheric heat), low reflectance, little evaporative heat loss, and rapid heat transmission. In urban areas the heat emission is higher than in rural areas. Air pollution in urban areas is very high, influencing the creation of the greenhouse effect. On hot summer days, the city feels the heat waves released by the dark-colored streets and buildings. Before night falls, the roads are still emitting heat, at a time when in the outskirts of the city, the rural areas are cooling down rapidly. Almost every city in the world nowadays is warmer than its surroundings. This temperature difference between urban and rural areas is called the "*urban heat*" effect. Urban heat is an effect that creates uncomfortable conditions, increases the need for air conditioning, creates great stress and what is most important negatively affects the environment, specifically urban smog (eg in metropolises).

2.2. Thermal comfort

Thermal comfort is considered the desirable or positive state that a person experiences in relation to how warm or cold that person feels. So it is very much related to the environment where the person is. Thermal comfort affects not only comfort, but also productivity, health and well-being.

Data on the human body:

a) Heat generated by the human body: 100W – by a person sitting; up to 1000W - from a person with maximum physical exercise.
b) Human body temperature is
c) A person will feel uncomfortable or sick if the body temperature is higher.

The body generates a certain amount of heat due to the oxidation process of food and this can be dissipated if the body temperature is not too high. Conversely, if a large amount of heat is lost to the environment, then the body temperature will drop. The clothes that people use play the role of thermal insulation. The amount of clothing used influences the amount of heat lost and this is reflected in the person feeling cold or warm.

The heat exchange between the body and the surrounding environment is represented by these quantities:

a) Evaporative loss (evaporation and sweating of the skin) and breathing.
b) Convection loss.
c) Radiation loss/gain.
d) Conductivity loss (usually negligible).
e) Gain from metabolic heat production.

The feeling of thermal comfort depends on a number of personal factors such as age, gender and body condition. However, for a group of people there are only two personal factors that have a relationship with comfort, and they are: physical activity and clothing size.

2.3. Thermal protection of the building

The thermal protection of the building includes the measures that must be taken, in order to reduce thermal losses from the inside/outside during the winter, while preventing excessive heat gains (from solar radiation) during the summer. In order for the heat losses to be as small as possible, i.e. to ensure a better heating with as little expenditure of thermal energy, namely electricity, it is necessary to take protective measures, first of all by the designers (architect and the constructor), associated with the corresponding relation.

Thermal protection of the building affects:

A) Improvement of comfort during the stay at work and in general during the stay in the residential premises.

The thermal comfort of a room presupposes the invariance of the temperature of the air and surrounding surfaces with time, as well as the non-existence of air currents. The design of the building must meet this requirement. Thermal comfort means that there is harmony between the person and the environment that surrounds him, which depends on age, race, gender, the work he performs, etc.; and which varies to some extent from the climate of the country where he lives: the conditions of comfort established for one people do not apply to another people. Man spends most of his life indoors. If we do not respond to the hygienic conditions of stay / comfort (residence), then people's health can be damaged.

B) Reduction of energy costs for heating (in winter) and cooling (in summer).

Energy resources in the world are decreasing, while with the development of society, thermal and electrical energy costs are increasing. Consequently, efforts are being made in the world to find and use new, renewable, alternative energy sources. In parallel with this, the costs of thermal energy for heating are increasing. In our country, a lot of electricity is used for heating, since it is useless from a thermodynamic point of view.

The thermal protection of the building greatly improves the economic use of thermal and electrical energy, bringing about its saving. So, with the improvement of the thermal protection of the building, the thermal energy required for heating the building, both hourly and yearly (during the winter period – converted into fuel), is greatly reduced. Meanwhile, the costs for cooling the premises in the summer period are significantly reduced (electricity). The addition of expenses for the thermal insulation of building structures, for the reduction of thermal transmission (winter/summer), in relation to the current state of constructions, is relatively low, compared to the benefits achieved through thermal insulation (saving thermal energy / fuel / electricity). (OFFICE, 2009)

C) Reduction of the power of the heating plant, related investments and maintenance expenses.

Investments for the installation and maintenance of the heating / cooling plant constitute an important element that enters into the overall construction cost and is related to the thermal protection of the building.

Since various constructive elements with high thermal conductivity (for structural or architectural reasons) are integrated into the building structures, such as columns, concrete, etc. - the so-called "thermal bridges" -, in these heat-insulated, high-conductivity, colder areas, water droplets can appear on the inner side of the surrounding structure. The consequence of this wetting will be not only the reduction of the thermal resistance of the constructive element (i.e. the increase of thermal losses / heat transmission), but also the gradual cooling of its surface, which leads to the occurrence of the phenomenon of condensation of steam indoor air (presence of moisture).

The existence of the thermal bridge is understood by the eye, by a pronounced blackening of the inner surface in the corresponding area, the appearance of the phenomenon of condensation, the formation of a thin layer of mold. This phenomenon, which is associated with the wetting of the external surface of the building structure on non-thermally insulated perimeter walls (Glaser phenomenon), causes the fall of external and internal plaster, the crumbling of bricks, the weakening of concrete, the corrosion of steel reinforcement, etc. While the thermal protection of the building does not create moisture, there is no deformation of the floor and other wooden constructions (placed on the walls), the color of the plaster is not removed, etc. The maintenance services of a thermally protected building are significantly lower and the life of the building longer.

2.4. Design requirements

The design requirements are presented in the necessary documentation of the building, in the plans and sections, orientation, height of the location where the construction takes place, construction structures, climatic conditions of the country, fuel to be used, etc. Architecturally (from a thermal point of view) more closed, compact forms should be preferred against open ones that have a larger external surface, thus greater thermal losses.

Calculations of the amount of heat are divided into two parts:

a) One concerns the transmission of heat from the various surrounding structures of the building, such as: walls, ceilings, floors, windows, doors and thermal bridges.
b) The other with the heat needed to warm the infiltrating outdoor air to room temperature.

Heat transmission is calculated from the known characteristics of the building materials, while air infiltration (the amount of air exchanged) is determined by experience. This amount of exchanged air constitutes natural ventilation, which is realized thanks to the change of internal and external conditions. The calculation of the necessary energy, either in a general way, in the first phase, or detailed, in the later stages of the design, requires the availability of climatic data for:

a) Calculation of cooling and heating requirements.
b) Design of heating, ventilation and air conditioning (HVAC) systems.
c) Energy assessment of buildings.

To calculate the energy needed for heating, you must:

- The internal temperature of the premises.
- External design temperature.
- Internal thermal sources (household appliances and solar radiation, which is generally favorable for heating, while unfavorable for cooling in summer).
- Phenomena that can cause cooling of the building, such as, for example, cold local wind.

The structure of the surrounding elements of the buildings must be calculated and dimensioned from the hygrothermal point of view, so that through the operation of the heating installations, the thermal comfort of the residence is achieved (in civil buildings), or the microclimate conditions in the case of technological processes (in industrial buildings).

For the surrounding elements of constructions, the following are defined in advance:

- The minimum thermal resistance is necessary to limit the thermal flow and to avoid the condensation of water vapor on the inner surface of each element.
- The thermal stability necessary to limit the internal surface temperature fluctuations of the structural elements.
- Resistance to the penetration of water vapors of constructive elements, which prevents the condensation of vapors inside the respective elements.
- Resistance to air infiltration.

With the orientation of the building from the south, the heat transmission has the most suitable values for heating the premises in winter and cooling them in the summer, where less heat is introduced by radiation, compared to the east / west orientation. The largest external surfaces of the premises should be oriented to the south; especially those environments where people stay more. Wind direction is also important, especially in winter. Strong winds (high speed) increase heat losses, especially in isolated buildings (open areas), or on the upper floors of taller buildings, compared to neighboring buildings. For this reason, the building should be placed so that it is as little exposed to the wind as possible, and preferably facing the wind are the parts of the building with fewer windows and good kissing (well kissing windows and full walls).

The rooms that are heated must be placed in the environment protected from negative external factors, i.e. in the interior of the building, while the rooms that are not heated, or secondary, such as the stairs (which, in multi-story buildings, in conditions of lack of electricity and to save it, it is good to provide windows, suitable for natural lighting).

Multi-story buildings consume less heat than single-story buildings, or with fewer floors, with the same usable area. As the number of floors increases, the percentage of thermal losses from the floor and terrace decreases, while that from walls and windows increases.

The constructive form of the building is important for thermal transmission (in thermal losses in winter and heat inputs in summer). For the same surface area, square and circular shapes have the smallest thermal losses (walls, windows). Buildings with large window surfaces, especially those with windows, are large consumers of heat in the winter, while their protection from radiation in the summer is quite difficult, and the consumption of cooling energy is high. Therefore, it is very important to insulate the walls as well as possible and to take measures to reduce the losses from the windows. The improvement of hygrothermal properties of perimeter walls, floors and terraces is achieved by increasing their thickness and thermal insulation, i.e. increasing thermal resistance, thermal stability, resistance to the penetration of water vapor and resistance to air infiltration. In this way, the value of the heat transmission coefficient "k" is reduced, the temperature difference in the different layers and on their inner surface is increased (thereby the temperature difference between the ambient air and the surrounding inner surfaces is reduced, thus improving sensitive thermal comfort), and at the same time the negative effects from the condensation of water vapor in the layers of the construction are avoided. We will thus have a higher internal surface temperature in winter and colder in summer. At the same time, some local measures should be taken; as, increasing the thermal resistance of the window parapet, where the radiator is placed, and the placement of heat reflective sheets there; increase in thermal resistance in thermal bridges in corners, etc. The most economically and thermally suitable material is cellular polystyrene.

Vertical building walls with 25 cm – 30 cm plastered on both sides have a heat transmission coefficient from 2.4 W/m^2K to 1.8 W/m^2K, which brings the heating need, referred to 1m2, from 100 – 120 W/m^2, and the amount needed for annual heating for $1m^2$ from 150-170 kWh/m^2. Whereas these loads today in European countries are respectively 30-40 W/m^2 and 50-70 kWh/m^2. These values correspond to an energetic building with thermally insulated walls with a coefficient of 0.35 – 0.45 W/m^2K (i.e. the thermal resistance of the walls should increase to values of 2.2 – 2.8 m^2K/W. Windows also play an important role. Therefore, the glass surface must be limited, within the values indicated by the thermal comfort; the thermal resistance of the glass must be increased to the values of 0.40 - 0.50 m2K/W, which can be achieved by using doppio insulating glass and in this way the overall heat losses are reduced by 20-40%. The windows must be large enough to provide sufficient natural lighting, as they have the lowest surface temperature and, therefore, create the greatest cooling of the human body through cold radiation in winter.

The indoor air in contact with the glass cools down a lot, increases its density and moves down towards the floor, creating a feeling of coldness at the feet. This effect is increased if the window is not well sealed. Another disadvantage of large windows is solar radiation in the summer, especially for windows facing east and west. External shutters greatly reduce the heat that enters the interior spaces through solar radiation. So, in an environment that is heated in winter, we have surfaces with different temperatures, some colder, such as external walls and windows, others warmer, such as stoves, radiators, etc. Thus, temperature differences appear in different points of the room, the size of which depends on the heating method. Through outdoor air and indoor air, we have temperature and humidity differences, which are accompanied by disturbing air currents. To avoid the occurrence of condensation, thermal protective measures should be taken in kitchens and bathrooms. For this reason, double-glazed windows should be provided, located so as not to cause the feeling of cold. The idea of building bathrooms without windows, with ventilation, is not recommended from a hygienic point of view.

Thermal comfort in moving air is particularly great, when all walls have an ambient air temperature of 20˚C. The term "walls of the environment" refers to all surfaces that give or receive radiant heat from the people staying in the building. The temperature of the walls of the building must be harmonized in such a way that heat losses, biologically necessary for the body, are not inhibited or increased (cooling).

3. Methodology

Facade typologies are different, but we broadly divide them into:

a) Opaque (complete) facades – are constructed of materials such as: masonry, stone, precast concrete panels, metal cladding, insulation, steel panels (cold formed).
b) Glass facades – are constructed of transparent glass materials and metal frames.

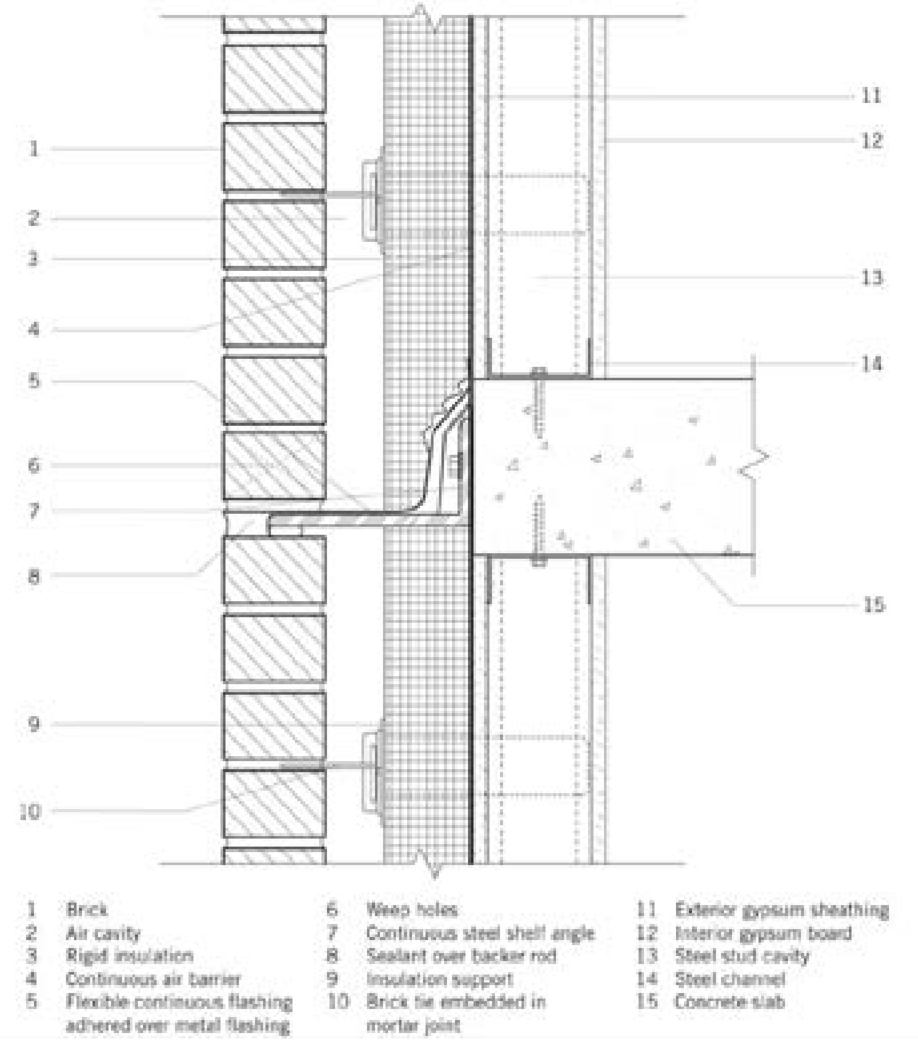
The physical behaviors of these facades are different and this is related to the components, materials and methods of their construction which are different.

In this article we will analyze only opaque facades and mainly its 3 main types. Below are their constituent elements and details of their installation in the elements of the supporting structures of the buildings.

3.1 Masonry facade

They consist of non-structural brick on the outside, supported by steel profiles (cold formed) or by a concrete masonry block (CMU). Between the outer layer of bricks and the internal support system there is an open space (cavity) which serves as a drainage area. The cavity makes it possible for water, which penetrates the outer layer, to drain at the bottom. Rigid insulation within the cavity or batt (glass wool) insulation between the steel frame members will improve the thermal performance of the wall. When the masonry cladding is supported by concrete masonry blocks (CMU), then insulation is placed on the outside of the CMU.

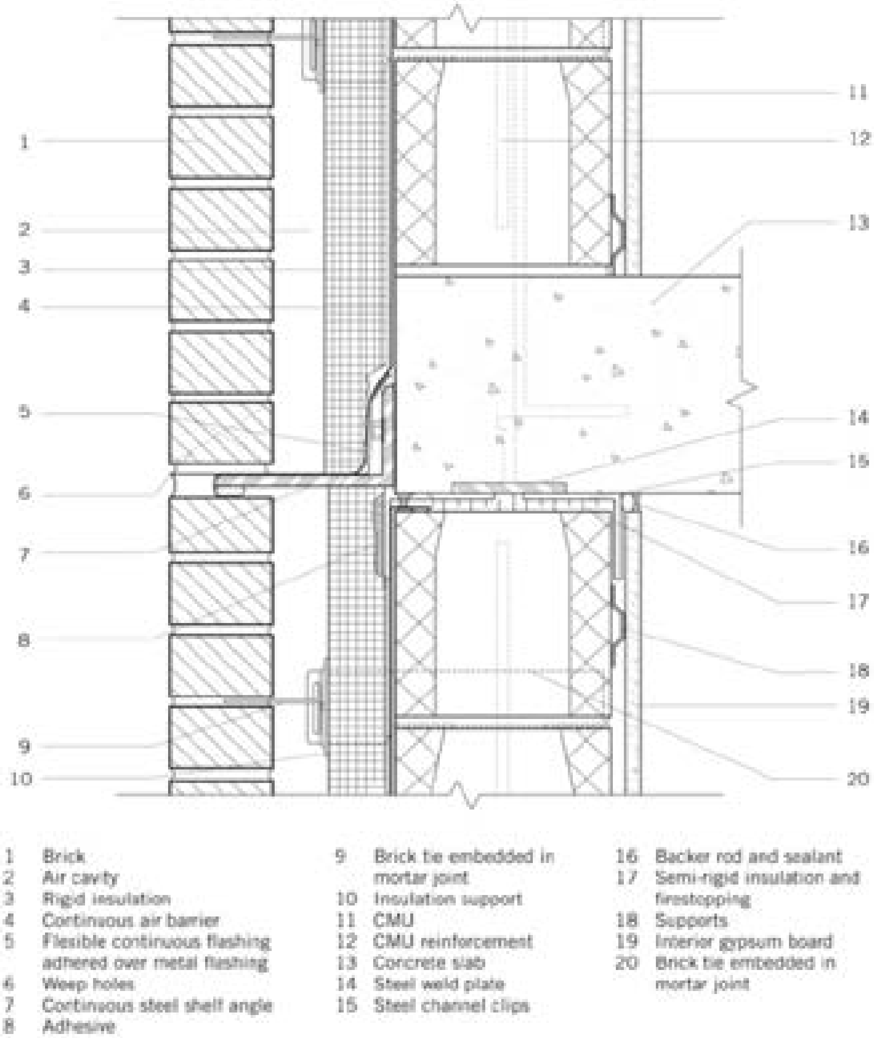
FIGURE 1 Masonry facade (Brick, hollow, frame, steel stud wall)



3.2. Facade with concrete panels

There are different types of concrete facades, such as: precast concrete panels used as a second structure cladding material, cast-in-place concrete walls, and insulated reinforced concrete panels. Other types of concrete facades include insulated concrete forms (ICF) and insulated concrete blocks (ICB). ICFs consist of pressed or expanded polystyrene panels that act as a frame for the concrete poured into the work; a final coat covers the polystyrene panels inside and out. These types of facades are typical for small-scale residential or commercial buildings. ICBs are CMUs that have expanded polystyrene sandwiched between the two faces of each block.

FIGURE 2 Masonry facade (Open brick wall with CMU support).



3.3. Rainscreen facade

They changed the traditional concept of protecting the interior space of the building from the penetration of rain and moisture. Most non-rainscreen facade systems rely on two protective layers. The first protective layer, the outermost surface of the facade, is the first protective barrier against air and moisture, designed to stop all air and water. The second layer of protection aims to stop small amounts of air and water vapor that can penetrate through the first layer of protection.

The rainscreen concept uses a variation of the outermost layer of the facade, which is not designed to be airtight and watertight. Instead, it acts as a rain barrier but relies on a weatherproof inner layer to block air and moisture penetration. Water penetration occurs because of air pressure changes that can be controlled by placing an air gap in the wall. Between the outer and inner layers there is a ventilation air space that moves water out. The void space (including the interior surface of the space) acts as the first layer of protection against air and water penetration. The clothing material is chosen first for its appearance. Cladding is usually panelized and can be made of a variety of materials, such as: stone, precast concrete, terra cotta (facade brick), cement composite, crystallized glass or metal. Because the inner layer is not visible in the finished construction, it is not designed for its visual qualities. Instead, it must be designed to withstand wind and seismic loads, to thermally and acoustically insulate the building, and to prevent air or water from entering the building.

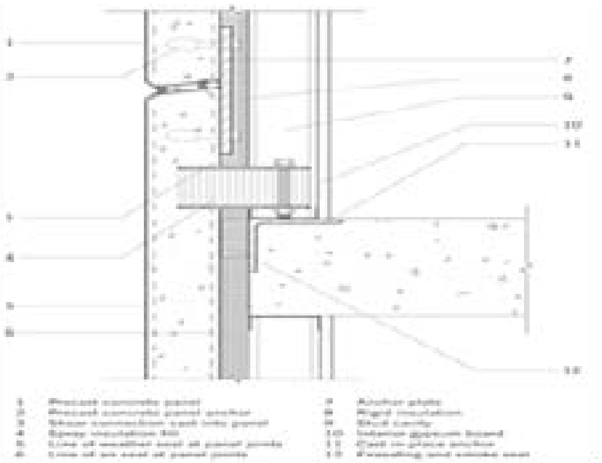


Figure 2 Precast concrete panel with steel

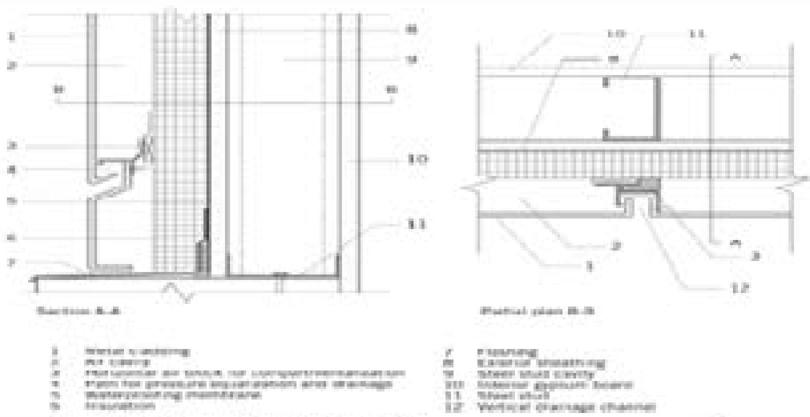


Figure 2 Rainscreen facade

4. Methods and analysis

4.1. Methods

Facade wrapping is very complex in itself and as such their calculation in terms of heat loss is a process that requires many influencing factors and parameters to be taken into consideration. Material selection is an important factor in designing sustainable facades. All materials have specific physical properties, such as density, thermal conductivity, thermal resistance and permeability.

R-values are most appropriate for determining the thermal performance of walls systems that consist of multiple layers of materials, each with its own R-value. Insulation materials are chosen for their thermal resistance and vapor barriers for their permeability. The nominal thermal resistance of a dark building facade can be calculated by adding the R-values of each layer of material, including air spaces. For each assembly, the overall R-value is calculated by adding the thermal resistance of the individual layers of material (determined by multiplying its R-values per unit thickness by the overall thickness of the material). Insulation materials for the three assemblies are rigid mineral insulation, spray polystyrene foam insulation, or mineral wool insulation.

4.2. Facade efficiency analysis

The components of the four facades taken into analysis are as follows:

A- ***Simple facade***: exterior plaster, thermal insulation, brick wall with holes, internal plaster.
B- ***Masonry facade***: masonry brick, air layer, thermal insulation; brick wall with holes, internal plaster.
C- ***Concrete facade***: coating foam-concrete panels, air layer, thermal insulation, brick wall with holes, internal plaster.
D- ***Rainscreen facade***: cladding aluminum panels, air layer, thermal insulation, brick wall with holes, internal plaster.

In summary, *Table 1* presents the part of the analysis of the facade decided for three of the types described above.

Table 1. Facade analysis

No.		Layers	δ (cm)	ρ (kg/m^3)	λ $(W/m\,K)$	$R = \delta/\lambda$ $(m^2\,K/W)$	$K = 1/R$ $(W/m^2\,K)$
1		Outside air				0.043	
2	*A*	*Simple facade with exterior plaster*	2	1800	0.90	0.022	45.00
	B	*Masonry facade*	*12*	*1800*	*0.60*	*0.200*	*5.00*
	C	*Coating foam-concrete panels*	*6*	*1600*	*0.70*	*0.086*	*11.67*
	D	*Cladding aluminum panels*	*1.4*	*2700*	*3.33*	*0.004*	*237.86*
3		Air layer	4	1.3		0.160	6.25
4		Thermal insulation layer	5	25	0.039	1.282	0.78
5		Brick wall with holes	20	1100	0.60	0.333	3.00
6		Internal plaster	2	1800	0.90	0.022	45.00
7		Indoor air				0.120	
		Total Value					
A		*Simple facade with exterior plaster*				***0.377***	***2.652***
B		*Masonry facade*				***1.998***	***0.501***
C		*Coating foam-concrete panels*				***1.883***	***0.531***
D		*Cladding aluminum panels*				***1.802***	***0.555***

5. Results

This article shows how the design of sustainable facades takes into account many factors, including environmental conditions, the orientation of the building, the design of the facade and the properties of its materials and components.

According to the analysis of different types of facades, it is noted that a ventilated facade has an energy efficiency of about 5 times more than a simple facade. While among the different types of ventilated facades taken in the study, we have similar values of thermal resistance, with a difference of 5% to 10%.

The physical behavior of the facade is a major contributor to the energy use of a building. Designers must consider the characteristics of materials and components, such as thermal and embodied energy, that go into the construction of a building's facade.

6. References

Aksamija, A. (2013), Sustainable Façades: design methods for high - performance building envelopes, New Jersey.

Aksamija, A., & Ghaffarianhoseini, A. (2018). The Role of Building Façade and Envelope Systems in Energy Performance of Buildings.

Balocco, C. (2002). A simple model to study ventilated facades energy performance. Energy and Buildings, 34(5), 469–475. https://doi.org/10.1016/S0378-7788(01)00130-X

De Gracia, A., Castell, A., Navarro, L., Oró, E., & Cabeza, L. F. (2013). Numerical modelling of ventilated facades: A review. Renewable and Sustainable Energy Reviews, 22, 539–549. https://doi.org/10.1016/j.rser.2013.02.029.

Lu, Y., Zhang, D., & Wang, Z. (2018). Thermal and Energy Performance of a New High-Performance Ventilated Façade System with Integrated Solar Collectors.

Shakti, R. & Huang, Y. (2020). Impact of Different Facade Types on Energy Consumption of Residential Buildings.

Shtjefni, A. (2014), Termodinamikë (Tabela dhe diagrama e vetive termofizike për disa lëndë), Tiranë.

Shtjefni, A., Akami, M. (2021). Termodinamika e aplikuar, Tiranë.

Vashtina, V., Shtjefni, A. (2009), Termodinamika, ngrohja dhe ajri i kondicionuar (fizika teknike), Tiranë.

Yaman, M. (2021). Different Façade Types and Building Integration in Energy Efficient Building Design Strategies. International Journal of Built Environment and Sustainability, 8(2):49-61. DOI: 10.11113/ijbes.v8.n2.732.

Applied Innovations in Industrial Robotics: Increasing Operational Range through Structural Modifications and Control System Analysis

MSc. Orges BESHKU[1]
IV-Automation Management Consulting sh.p.k

Abstract

This article focuses on the studying and improving of the capabilities of a Fanuc M-900iB, 6 degree of freedom industrial robot arm. The main objective is to increase its reach and expand its workspace in situations where parallel production lines in automated industries can use one robot with a wide work frame instead of two robots per each line. In order to achieve these objectives, a complete engineering study of the structural, mechanical, electrical, electronic and software components has been carried out. To make this analysis, evaluation and improvement, the maintenance and construction manuals of this robot, the computer program used in real life as well as its physical study and computer simulation were used as reference.

[1] Orges Beshku is a highly accomplished automation expert and technology innovator, with expertise in automation, robotics, and marine electronics. He is the Founder & CEO of IVy Ltd., where he leads digital transformation and automation management consulting projects across various industries, specializing in advanced technologies like IoT and Industry 4.0. Orges is also a Lecturer of Robotics & Electronics at the European University of Tirana, where he contributes to both teaching and research in the field. Currently serving as Head of Engineering at Durrës Port Authority Inc., Orges oversees major engineering initiatives, including the modernization of port infrastructure and machinery. Orges holds a Master of Science in Mechatronics Engineering from the Polytechnic University of Tirana, where his research focused on improving industrial robotic systems. He is certified in multiple technical and management fields, including PLC programming, advanced robotics, marine electronics, port management, and leadership in times of change.

In the mechanical aspect, the structure is studied with all the constituent components such as the base, the joints, the end effector and the links between them with the weight of each component, electrical consumption, ranges of motion, angular velocities, moments of inertia as well as internal components have been considered while the second part of the mechanics covers the kinematics and dynamics in mathematical aspects and computer simulations of the trajectory of its movement, considering the inertia created in the minimum and maximum capacity.

In the electrical/electronic aspect, the configuration has been studied in detail including the control unit, main board, servo amplifier, servo motors, power supply unit and transformer. In the second part, the electronic control method, general control block diagrams as well as specific position, speed, and current control block diagrams are presented, explaining the filters and amplifiers associated with each, Bode diagrams, and characteristics of open and closed systems. In conclusion links 1 and 3 were extended by 0.3m and 0.25m, offering us a much wider and effective working frame of the robot in the workplace. According to the simulations performed the forces required by the change are affordable and within the safety factor for the supply unit with power, amplifier and servo motors.

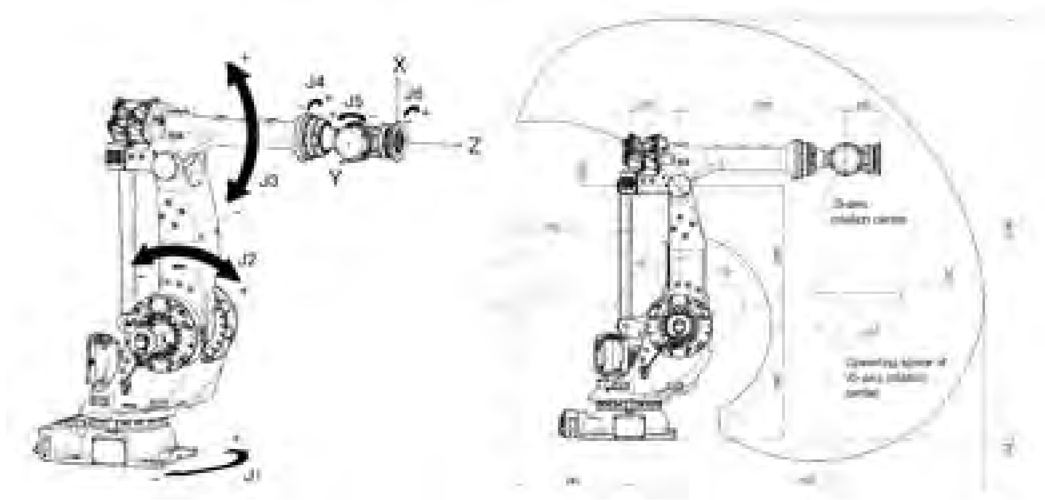
Keywords: *Industrial Robotics, Control System Analysis, Structural Modification, Automated Industries, Structural & System Integrity, Kinematics & Dynamics*

1. The robotic system

The field of robotics has progressed rapidly over the past few decades, with an increasing number of robots being used in various industrial and commercial applications. One of the most popular and versatile robots used in industry today is the FANUC M-900iB. This is an industrial robot designed for applications that require high speed and load capacity with a maximum capacity of 360 kg and a reach of 2655 mm. The six-angle-of-freedom configuration of the robot allows for a wide range of motion and flexibility in a variety of applications.

In this research article, the dynamics and control of the FANUC M-900iB/360 robot, the various factors that contribute to the robot's performance, including its mechanical structure, actuation system and control algorithms, will be examined in detail. By analysing these factors, we aim to gain a deeper understanding of robot behaviour and dynamic control and provide insights that can lead to further improvements in the performance and capabilities of this important industrial tool.

FIGURE 1 – Mechanical Structure



1.1 AI Generated scenarios

Two parallel production lines with two Fanuc robots:
In this scenario, a factory operates two parallel production lines, each producing boxes of varying sizes and colours. Two Fanuc robots, one positioned beside each production line, are tasked with palletizing the boxes onto adjacent pallets.

One central robot with elongated arms handling two production lines:
In this scenario, a single Fanuc robot, centrally located between two parallel production lines, is responsible for palletizing the output from both lines. The robot has elongated arms, allowing it to reach the boxes coming from both lines. This setup reduces the need for multiple robots, optimizing space and operational costs while maintaining efficiency.

FIGURE 2 – Standard M900-iB robotic arms working on 2 production lines

FIGURE 3 – Single modified M900-iB robotic arm working on 2 production lines

2. Methodology

The robot manufacturer's own manuals, descriptions, specifications and methods will be used to analyse and evaluate the robotic system in accordance with the robot's function. The analysis will be done to work as a robotic palletizing arm and will be compared with the real work values of the robot to prove the accuracy of the simulation. Also, mathematical models assisted by simulations in software like Matlab, RoboAnalyzer will be used to evaluate the capabilities of the current system and conclude if this system is able to withstand structural changes based on the changes made.

3. Mechanics

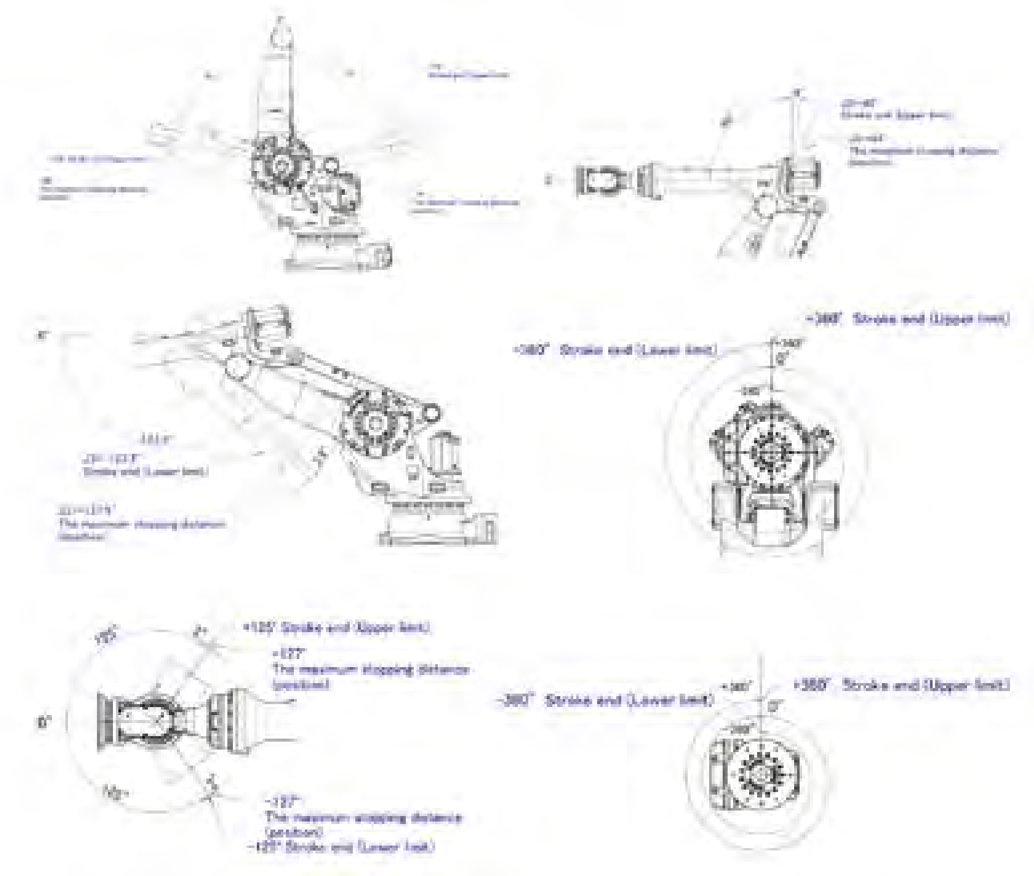
3.1 Mechanical Structure

The mechanical structure of the FANUC M-900iB/360 robot consists of several key components, including the base, six joints (J1-J6), links, and end effector.

The base of the robot is usually mounted on a pedestal or a floor fixture and provides a stable base for the rest of the robot. The six robot joints (J1-J6) are responsible for providing the robot with six degrees of freedom, allowing it to move in different directions and orientations. The end effector of the robot is responsible for performing specific tasks, such as grasping or manipulating objects.

Overall, the mechanical structure of the FANUC M-900iB/360 robot is designed to provide maximum flexibility and manoeuvrability while maintaining stability and precision during operation.

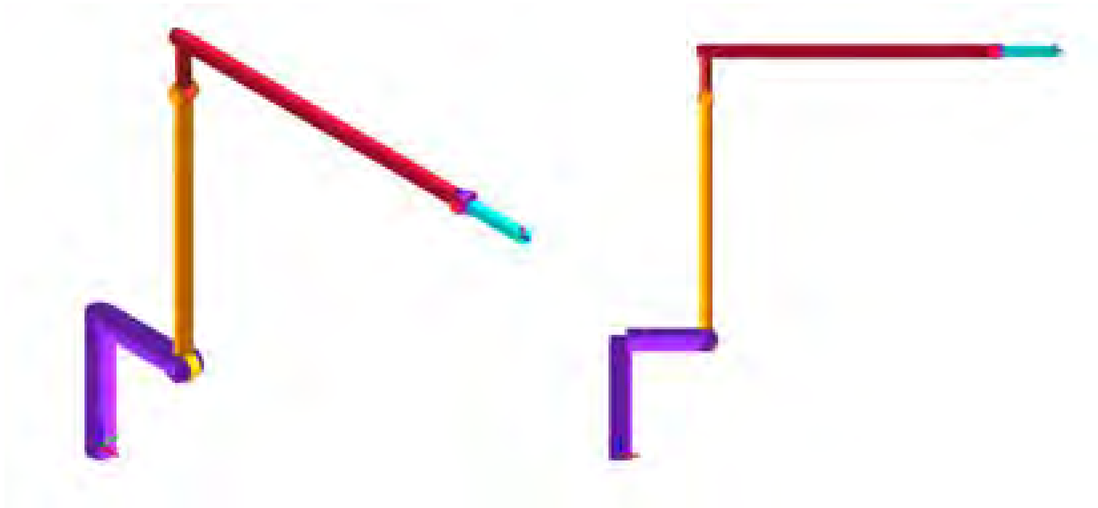
FIGURE 4 – Degrees of freedom and their limits



3.2 Kinematics

Forward kinematics is the study of the manipulator to find the position and orientation of the tip or end effector using the joint values of the manipulator. The first step of performing forward kinematics is determining the lengths of the links, the second step is setting the frames and finally, through the obtained matrices, we find the final homogeneous matrix.

FIGURE 5 - FANUC M-900iB simulated on RoboAnalyzer



D-H parameters can be found based on frame assignment. These parameters are illustrated in table 4:

FIGURE 6 - D-H Parameters

Link no.	Link type	Link length d (z) (m)	θ (°)	Link displacement a (x) (m)	α (°)	Initial angle	Final angle
1	R	0.5	Var	0.37	90	0	θ_1
2	R	0	Var	1.05	0	0	θ_2
3	R	0	Var	0.2	90	0	θ_3
4	R	1.25	Var	0	90	0	θ_4
5	R	0	Var	0	-90	0	θ_5
6	R	0.26	Var	0	0	0	θ_6

The transformation matrix for a link i is described as follows:

$$A_i = \begin{bmatrix} \cos\theta i & -\sin\theta i\cos\alpha i & \sin\theta i\sin\alpha i & ai\cos\theta i \\ \sin\theta i & \cos\theta i\cos\alpha i & -\cos\theta i\sin\alpha i & ai\sin\theta i \\ 0 & \sin\alpha i & \cos\alpha i & di \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

And the resulting matrices for each link are:

$$T_1^0 = \begin{bmatrix} \cos\theta 1 & 0 & \sin\theta 1 & 0.37\cos\theta 1 \\ \sin\theta 1 & 0 & -\cos\theta 1 & 0.37\sin\theta 1 \\ 0 & 1 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_2^1 = \begin{bmatrix} \cos\theta 2 & -\sin\theta 2 & 0 & 1.05\cos\theta 2 \\ \sin\theta 2 & \cos\theta 2 & 0 & 1.05\sin\theta 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_3^2 = \begin{bmatrix} \cos\theta 3 & 0 & \sin\theta 3 & 0.2\cos\theta 3 \\ \sin\theta 3 & 0 & -\cos\theta 3 & 0.2\sin\theta 3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$T_4^3 = \begin{bmatrix} \cos\theta 4 & 0 & -\sin\theta 4 & 0 \\ \sin\theta 4 & 0 & \cos\theta 4 & 0 \\ 0 & -1 & 0 & 1.25 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_5^4 = \begin{bmatrix} \cos\theta 5 & 0 & \sin\theta 5 & 0 \\ \sin\theta 5 & 0 & -\cos\theta 5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad T_6^5 = \begin{bmatrix} \cos\theta 6 & -\sin\theta 6 & 0 & 0 \\ \sin\theta 6 & \cos\theta 6 & 0 & 0 \\ 0 & 0 & 1 & 0.26 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In the base position of the robot where all the initial angles are 0°, we determine the positions and rotations between the base and the links as well as between the preceding links. The resulting matrices are:

$$\boldsymbol{T_0^1} = \begin{pmatrix} 1 & 0 & 0 & 0.37 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{T_1^2} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1.05 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{T_2^3} = \begin{pmatrix} 1 & 0 & 0 & 0.2 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{T_3^4} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1.25 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{T_4^5} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{T_5^6} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.26 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The resulting matrix gained from matrix multiplication shows the final rotation and position of the end effector of the robot.

$$T_6^0 = T_1^0 \times T_2^1 \times T_3^2 \times T_4^3 \times T_5^4 \times T_6^5,$$

$$T_6^0 = \begin{bmatrix} r11 & r12 & r13 & Px \\ r21 & r22 & r23 & Py \\ r31 & r32 & r33 & Pz \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$T_6^0 = \begin{bmatrix} R_6^0 & P_6^0 \\ 0 & 1 \end{bmatrix}.$$

$$\boldsymbol{T_0^6} = \begin{pmatrix} 1 & 0 & 0 & 1.88 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1.75 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

3.3 Computer simulation of dynamics

Robot dynamics is a subfield of robotics that deals with the study of the motion and forces of robots. It focuses on understanding the relationships between the motion of a robot, the forces that cause that motion, and the physical properties of the robot itself, and is important for their design and control. It is used to analyse the performance of these systems, including speed, accuracy and stability of movements.

To model and analyse the dynamics of the robot, mathematical models are used, such as the equations of motion, which describe the movement of the robot under the influence of forces. These models take into account the kinematics (position, velocity and acceleration) of the robot, as well as the forces and torques acting on it, but it is possible to simulate their dynamics in software and that is how the robot is. studied in this project. By calculating the dimensions, weight and moments of inertia of each link, a computer model identical to the physical one was designed to which a maximum load was applied to the end effector and its behaviour in movement with a complex trajectory was analysed.

This simulation not only gives us details of the movement of the robot with the data set by us, but also correctly shows the outer limits and the working range of the robot which coincides with that presented in the technical documents of the manufacturing company, which proves the accuracy of simulation.

As seen in figure 5 and 6, there are 2 tables where the first one contains the parameters of the original robot and the second one contains the parameters of the proposed robot, where what is noticed is the addition of 0.3m (from 0.5m to 0.8m) in the link first as well as for the third link with 0.25m (from 1.25m to 1.5m). Based on these values, the simulation and comparison between the systems will be done.

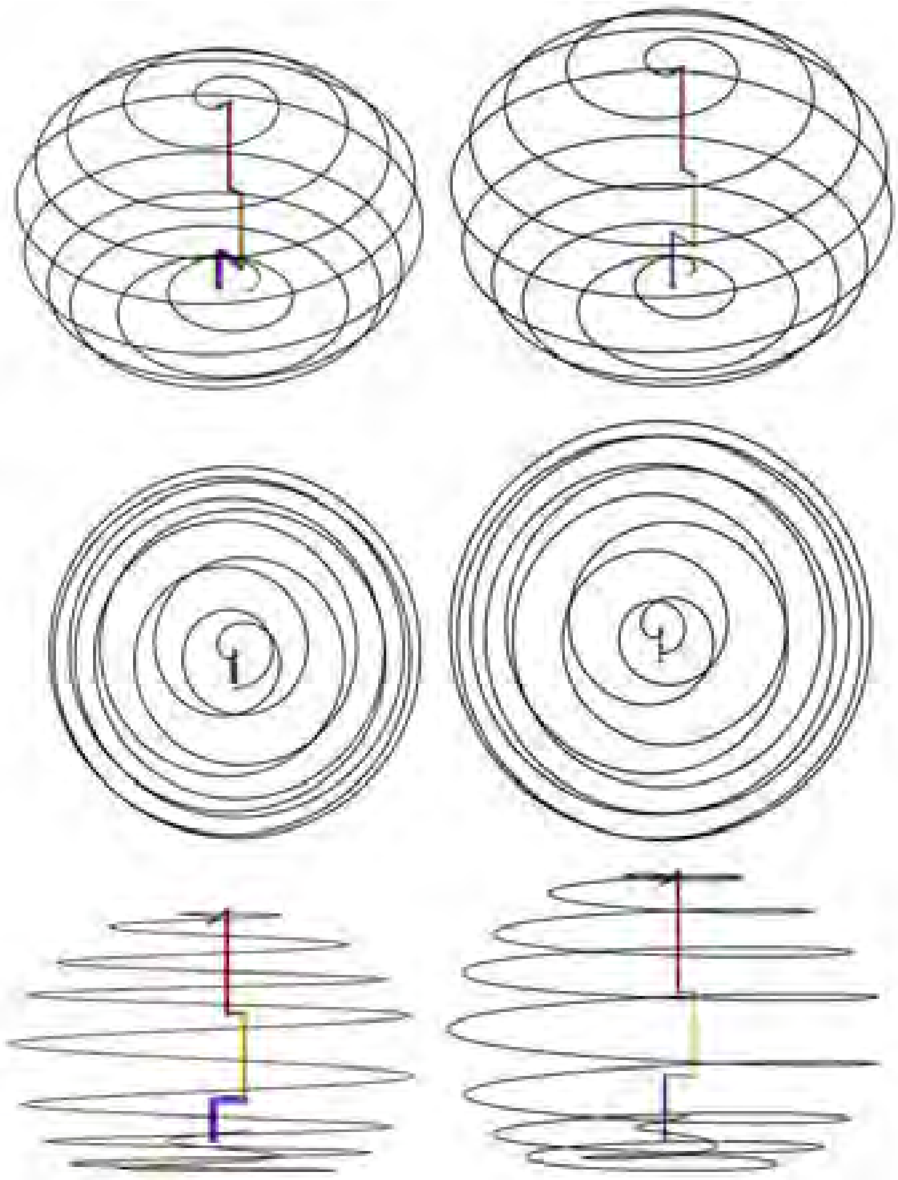
FIGURE 7 - D-H Parameters during simulation for original lengths

Link no.	Link type	Link length d (z) (m)	θ (°)	Link displacement a (x) (m)	α (°)	Initial angle	Final angle
1	R	0.5	Var	0.37	90	0	3600
2	R	0	Var	1.05	0	14	90
3	R	0	Var	0.2	90	-70	90
4	R	1.25	Var	0	90	0	720
5	R	0	Var	0	-90	125	-125
6	R	0.26	Var	0	0	0	720

FIGURE 8 - D-H Parameters during simulation for modified lengths

Link no.	Link type	Link length d (z) (m)	θ (°)	Link displacement a (x) (m)	α (°)	Initial angle	Final angle
1	R	0.8	Var	0.37	90	0	3600
2	R	0	Var	1.05	0	14	90
3	R	0	Var	0.2	90	-70	90
4	R	1.5	Var	0	90	0	720
5	R	0	Var	0	-90	125	-125
6	R	0.26	Var	0	0	0	720

FIGURE 9 - Simulated working range and frame before and after modifications.



As can be seen in Figure 7, the working frame range of the proposed robot is significantly larger than the working frame range of the existing robot. This is a positive result in the first simulations, but it should be taken into account that the extension of the links leads to an increase in the moment of rotation of the arm, bringing the need for more force. If the required force is within the parameters of the existing servo motors, then no other changes are needed to improve this system structurally.

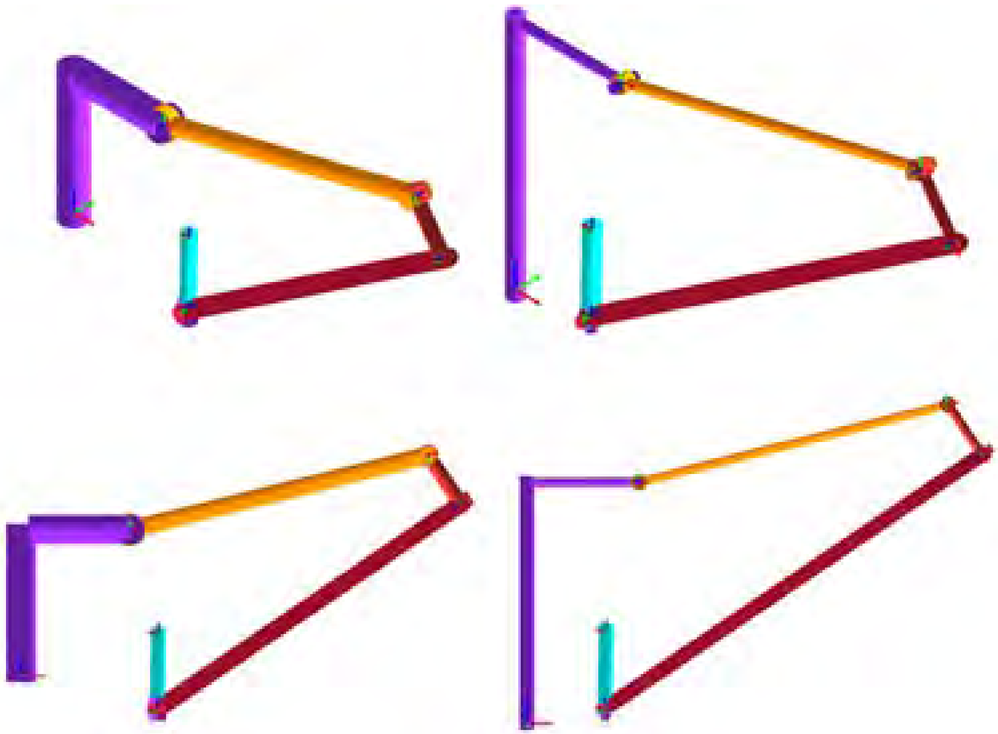
The values recorded in the program to correctly determine the forces are:

Link	Mass (kg)	Inertia I_{xx} (kgm^2)	Inertia I_{yy} (kgm^2)	Inertia I_{zz} (kgm^2)
1	350	16.6	23.26	25.28
2	280	27.85	4.2	27.85
3	240	21.8	21.25	3.05
4	60	0.76	0.76	0.6
5	20	0.26	0.21	0.25
6	20 + 360	0.1	0.1	0.1

In the created simulation, the geometric movements, speeds, accelerations and forces needed to perform the movements of the robot in the defined trajectory were measured. Each of the measurements is made at the endpoint of the link and each link is dependent on the previous one. The physical parameters are defined in the table above.

The starting point is set as figure 8 to ensure that during rotations we also get the maximum working range of the robot. Also in this figure is the simulation of the starting position with the proposed changes to the body of the robot.

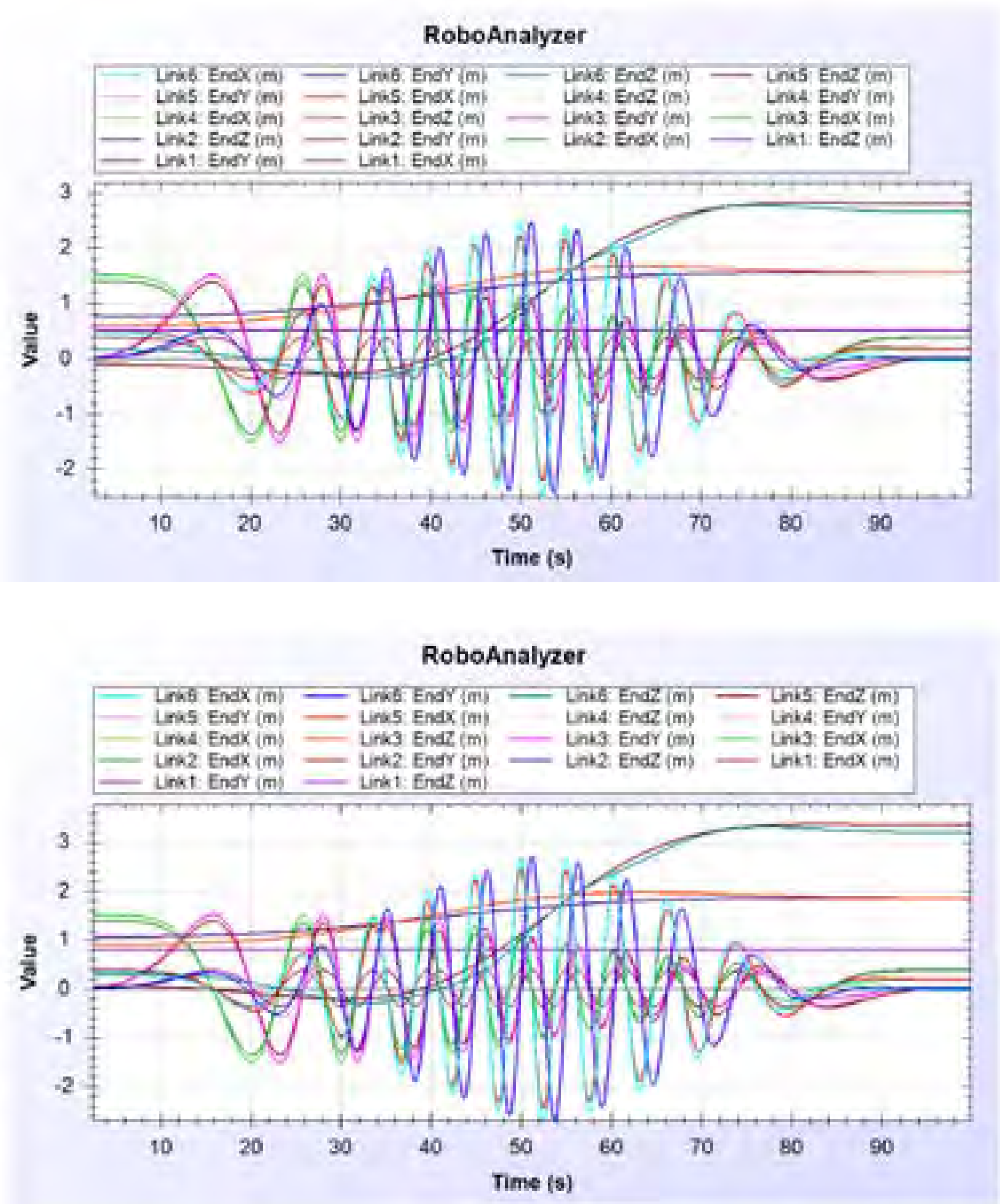
FIGURE 10 - Simulation of starting point position



In the summary graph (figure 9) we see and compare more clearly the trends of all links and their maximum values. What interests us and proves the accuracy of the geometric simulation is the maximum position of link 6 in the y and z axis. According to the specifications of the robot from the manufacturing company, the maximum values of y and z are 2655mm and 3160mm respectively, but in z the base of 350mm is also taken into account, which is not taken into account in the simulation of our robot. This makes the z of this a value of 2810mm.

The simulation shows that the maximum values of y and z of link 6 are 2655mm and 2810mm, respectively, which proves the accuracy of the simulation from the geometric point of view. While in the new proposed system we have a maximum z height of 3360mm and a maximum width of 2900mm, in contrast to the manufacturer's 2655.

FIGURE 11 - Simulation of the position of 6 links before and after modification



The graphs below (figure 12, 13 & 14) show the velocities, accelerations and forces of all links summarized to make a visual comparison. Whereas it seems for speed and acceleration link 1 plays the main role while for required force link 2 prevails.

FIGURE 12 - Velocities of 6 simulated links

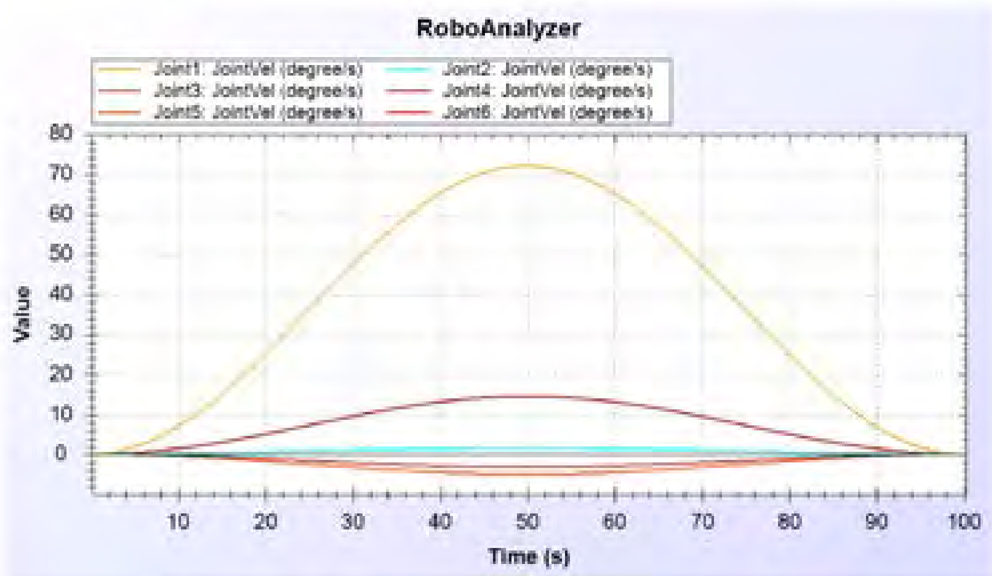


FIGURE 13 - Accelerations of 6 simulated links

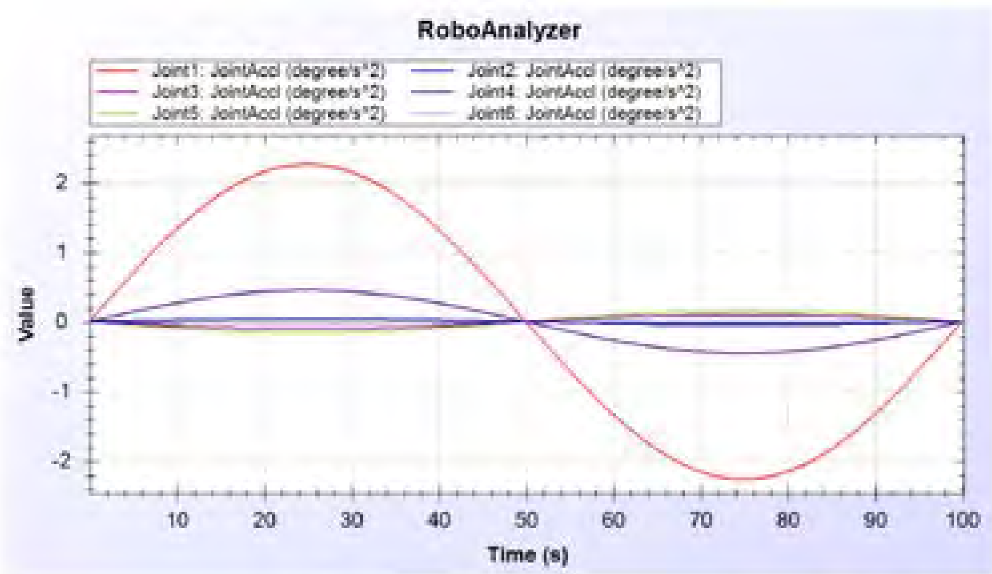
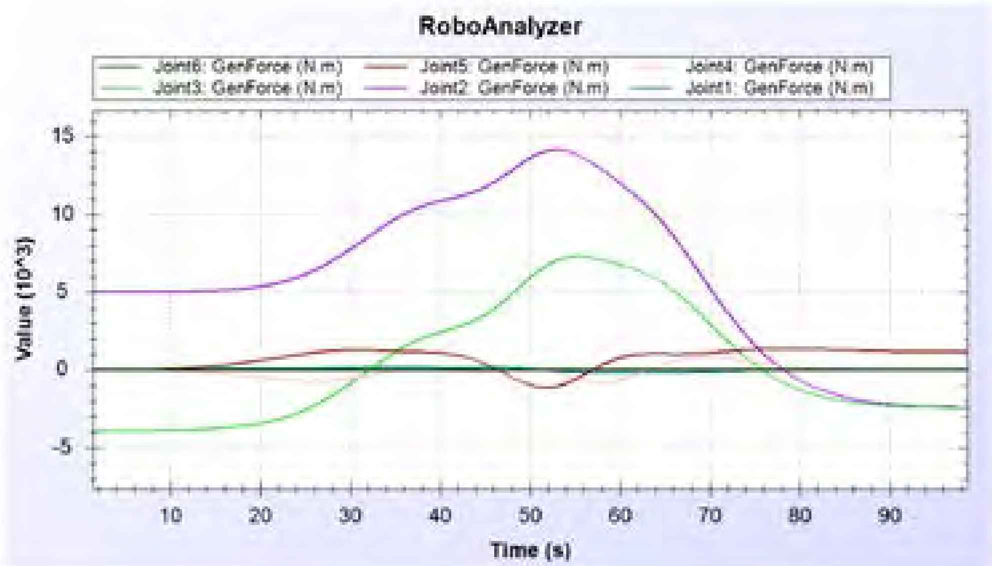
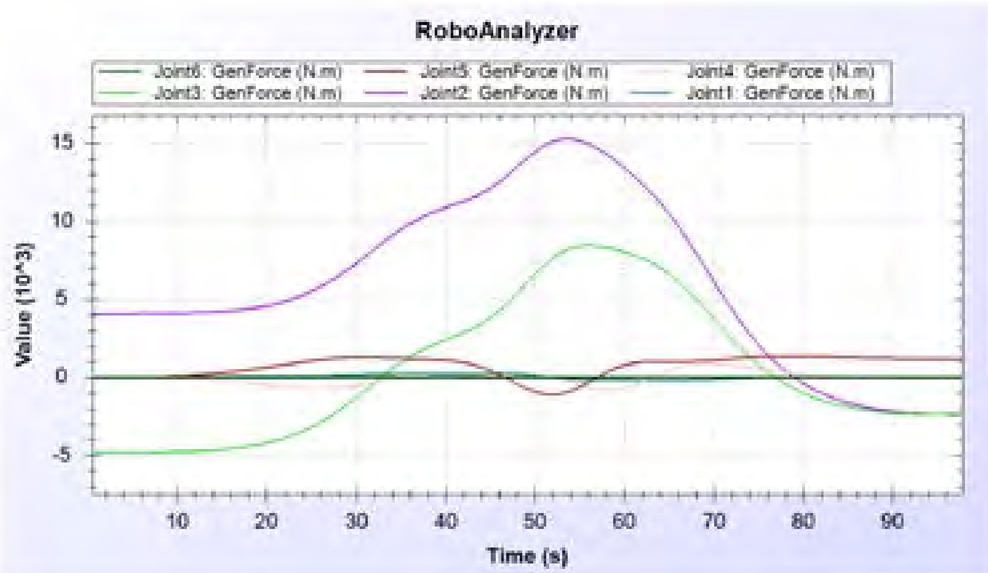


FIGURE 14 - Forces of 6 simulated links before and after modification





The maximum force that the links can withstand at maximum speed with maximum load, taking into account the inertia, is the total mass of the weight carried by the robot, which is:

360kg + weight of link 6, 20kg + weight of link 5, 20kg + weight of link 4, 60kg + weight of link 3, 240kg + weight of link 2, 280kg = 980kg.

From which the maximum possible moment is calculated 980kg x 9.805m/s2 = 9,609N x 2.56m (arm) = 24,600Nm

This does not account for the weight of the first link as it rotates around the z-axis and does not raise or lower concrete weights even though it carries the total weight of the robot as a quasi-static structure.

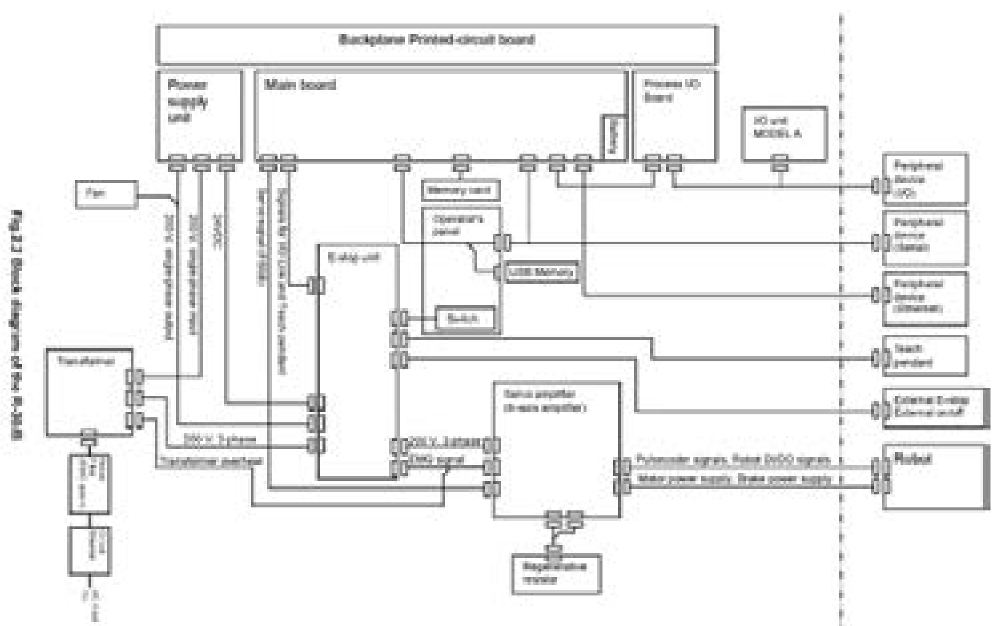
The second link has a total moment of 14,110Nm for the manufacturer's structure and 15,252Nm for the proposed structure, these values are far from the 24,600Nm maximum moment of the robot.

This concludes that our modifications of the links are acceptable by the simulation and can withstand the new forces. This leaves the evaluation of the amplifiers and control systems to conclude that these modifications are acceptable in reality without making changes in power and amplification.

4. Electric & electronic configuration

The connection between the components is made according to the scheme in figure 13 where after the voltage input there is protection by means of electrical switches, noise filter and then input to the transformer which makes the conversion to AC 220V and distributes the current to the supply unit (which makes also the conversion to DC 24V), the emergency stop unit and the servo amplifier. Meanwhile, the signals in these components come from the main and secondary panels, which are finally connected to external physical units such as the robot, emergency buttons, controller, ethernet communication, etc.

FIGURE 15 - Connection diagram of the main components



The main board

The main board contains a microprocessor, its peripheral circuits, memory and operator panel control circuitry. The main CPU controls the positioning of the servo mechanism.

- I/O printed circuit board, FANUC MODEL-A I/O unit

Various types of printed circuit boards are available for applications including I/O processing board. FANUC MODEL-A input/output unit can also be installed. When used, different I/O types can be selected. These are connected to the FANUC I/O Link.

- E-stop unit

This unit controls the emergency stop system of the robot controller. It also has the user interface terminal terminals for, relevant safety signals, external on/off signals etc.

- Power supply unit

The power supply unit converts AC power to various levels of DC power.

- Backplane printed circuit

The various control printed circuit boards are mounted on the backplane printed circuit board.

- Controller

All operations including robot programming are performed with this unit. Controller status and data are shown on the liquid crystal display (LCD) on the pendant.

This is also where the peripheral components are connected as in the figure above.

- 6 axis servo amplifier

The servo amplifier controls the servo motor, encoder signal, brake control, travel limits.

- Operator panel

Buttons and LEDs on the operator panel are used to turn on the robot and indicate the status of the robot.

- Transformer

The supply voltage is converted to the required AC voltage for the controller (220V), by the transformer.

- Fan unit, heat exchanger

These components cool the inside of the controller.

- Circuit breaker

If the electrical system in the controller malfunctions, or if the abnormal input power causes high current in the system, the input power is connected to the circuit breaker (switch/circuit breaker) to protect the equipment.

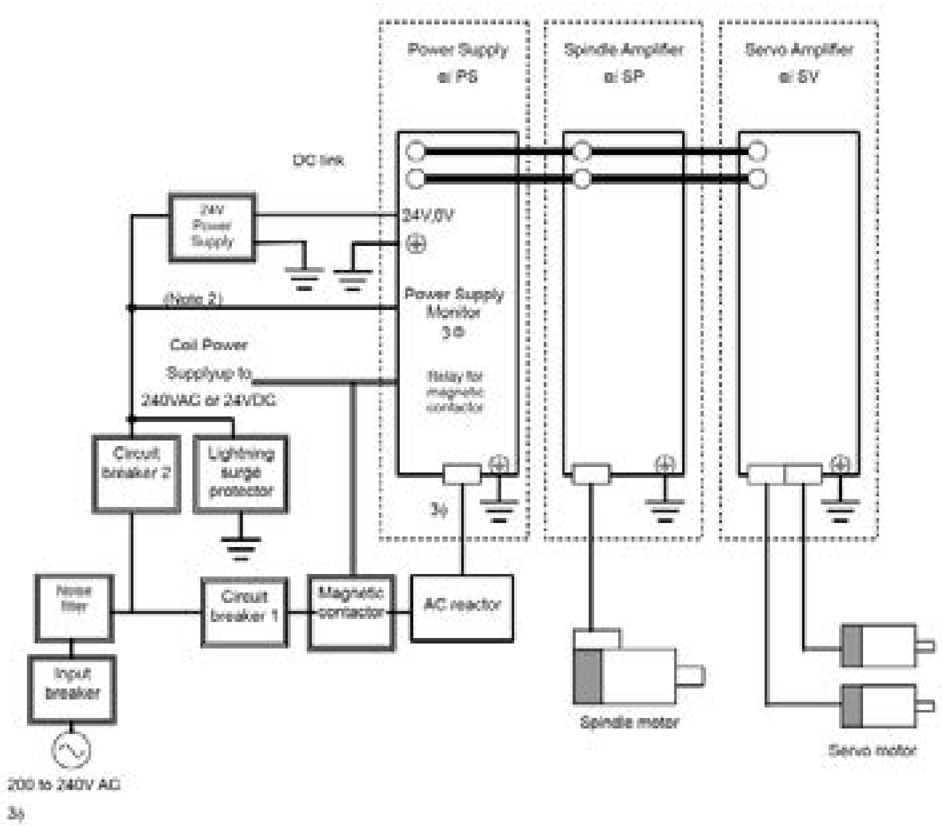
- Regenerative resistance

To discharge the resistive electromotive force from the servo motor, a regenerative resistor is connected to the servo amplifier.

4.1 Servo Amplifier

The servo amplifier is the unit that amplifies the output power that serves to drive the servo motors. As described in the previous chapter, the servo amplifier is supplied with 3-phase current from the emergency stop unit, also at the input are the emergency signals from this unit, the overheating signal from the transformer as well as the command input signals from the main circuit. at the output it is connected to the robot which supplies it with energy and exchanges input-output signals for the movements performed.

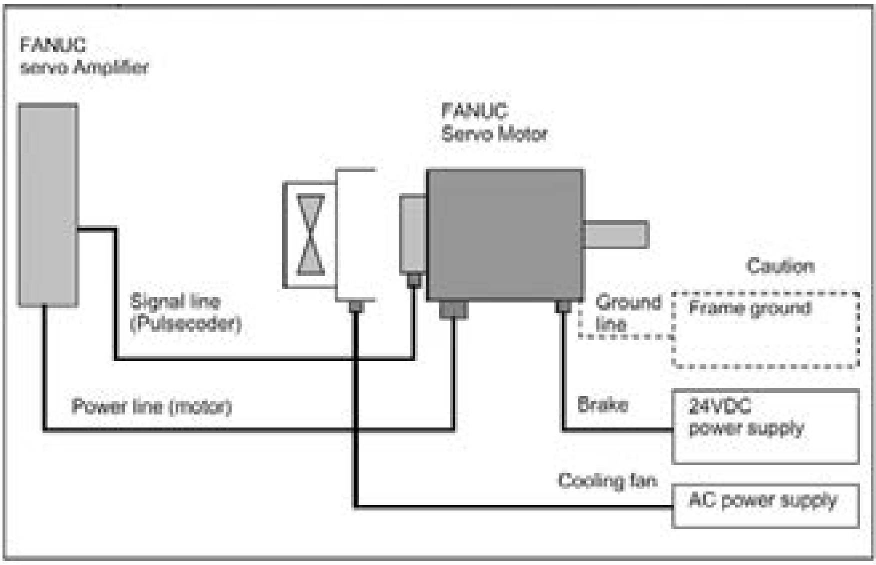
FIGURE 16 - Schematic of the power supply for the servo-amplifier



4.2 Servo Motor

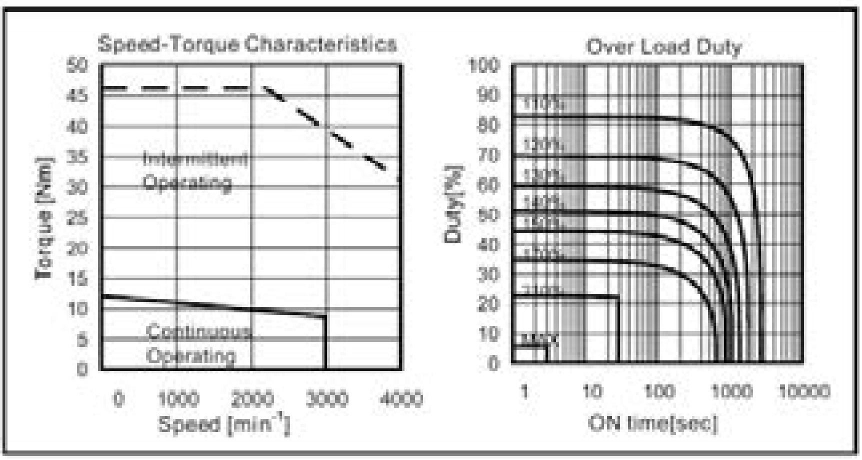
The internal circuit of the motor connection is as below (figure 14), where the power comes from the servo amplifier and by means of the pulse coder, the motor returns the signal to the amplifier to make the necessary speed adjustments. Meanwhile, the fan is also connected to the motor, which serves for cooling, grounding for discharging currents, as well as the brakes with DC 24V connection.

FIGURE 17 - Wiring diagram of the servo motor



In figure 16 we see the characteristics of the work according to the torque and in overload. The smallest model of this robot is taken as an example to create a general idea of the characteristics of these servo motors.

FIGURE 18 - Operating characteristics of the servo motor



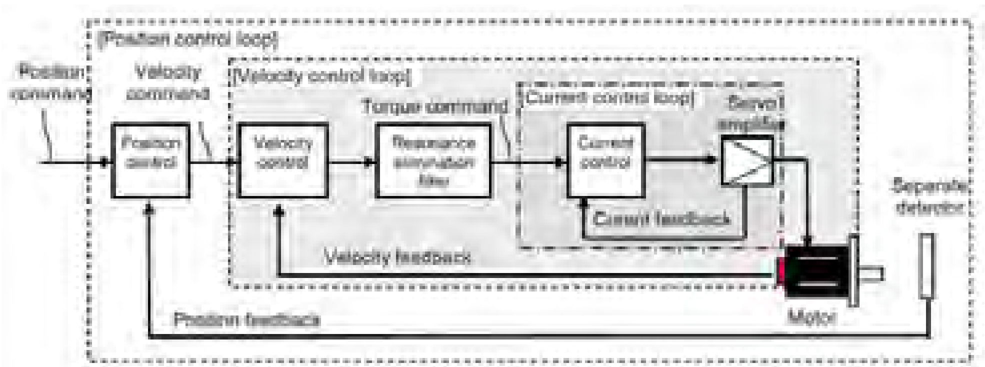
As can be understood from the graph, the maximum speed that the engine reaches in continuous operation is 3000 rpm, while in intermittent operation, but with a higher torque, it exceeds this rotation speed.

In the work with 110% overload, we see the maximum working interval is up to 3000s or 50min and the more the load increases the more the working time decreases and when 200% of the maximum load is reached the engine works only a few seconds before it burns out or fail.

4.3 Response control

In servo motor control, position, speed and torque (torque) are controlled by linking them together. Figure 17 shows a general engine control setup. Servo control is accomplished with this triple block diagram, which has a stream of successive speed changes needed to operate the arbitrary motion, speed command to do position control, torque command to do speed control as well as the subsequent changes in current required to operate at that torque to make current control.

FIGURE 19 - General scheme of control of a servomotor

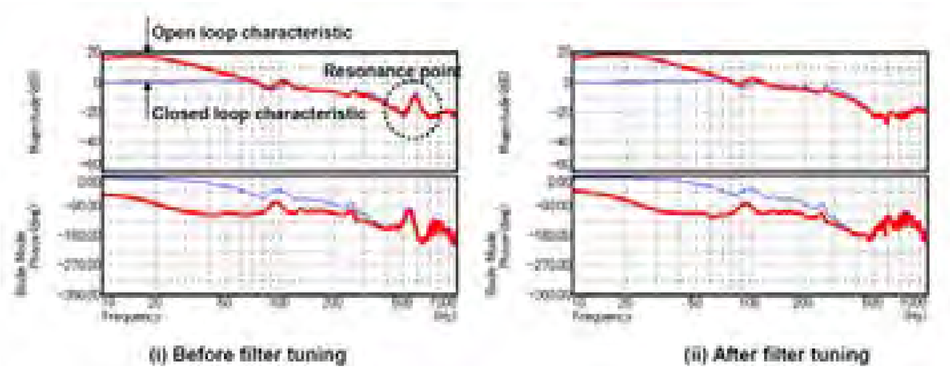


In short, the motor is controlled through individual control of position, speed and current.

In order to increase the tracking performance for position commands and to reduce the contour error, it is necessary to improve the response speed of the speed control and current control. Since each control loop is affected by the internal control loop, servo tuning must be done for the current control loop, speed control loop, and position control loop in the order mentioned in the Figure 17.

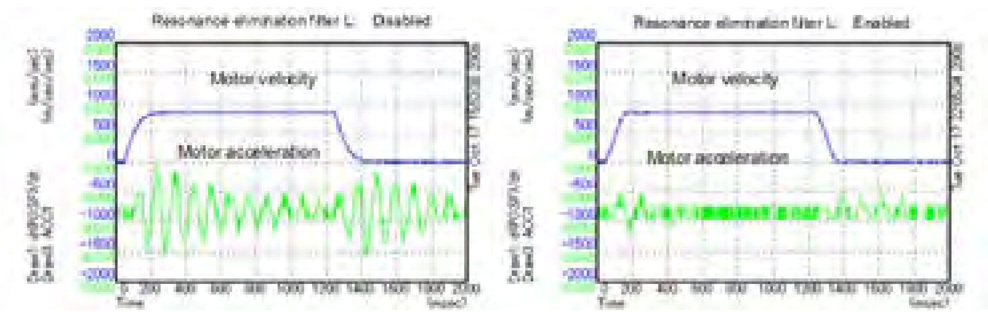
Figure 18 shows the resonances created by speed adjustments which are eliminated by the filters explained below.

FIGURE 20 - Schematic of filter resonance adjustment



The Resonance Elimination Filter function eliminates vibrations by applying a filter designed to eliminate components from a feedforward command/speed command.

FIGURE 21 - Resonance adjustment filter effects



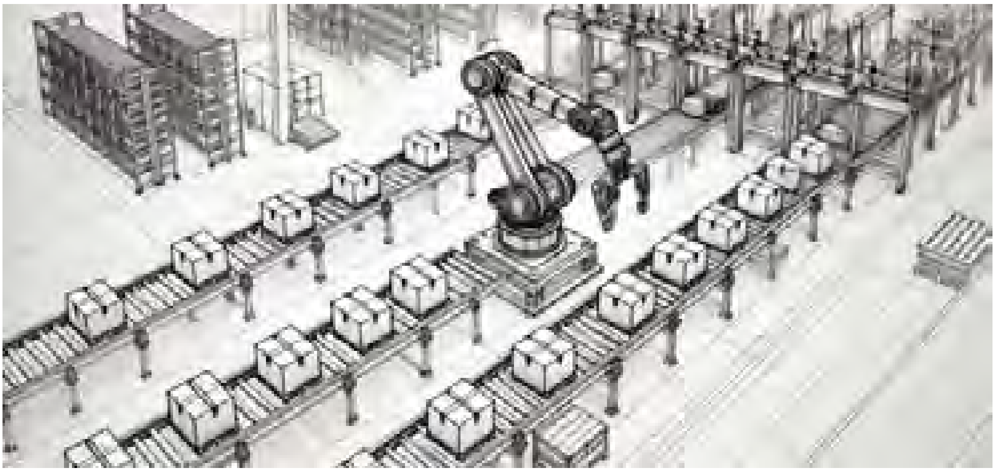
Conclusions

the study and analysis of the dynamics and control of the industrial robot with 6 degrees of freedom FANUC M-900iB/360 brought some valuable conclusions in the engineering nature that help us determine the functionality, capacities and limits of the use and manipulation of this robot.

Mechanical study:
In the mechanical aspect, the physical components of the robot were studied such as the base, links, joints, and end effector, as well as the second part of mechanics which covers topics such as forward kinematics in mathematical form, inverse kinematics and simulated dynamics of the robot.

Mechanical improvements:
Mechanical modification of links 1 and 3 by lengthening them respectively by 0.3m and 0.25m thus increasing the range of the robot's working frame without compromising the electrical and electronic integrity of the system, supported by computer simulations.

FIGURE 22 – AI made final scenario of modified robot between production lines



Electrical/electronic study:
In the electrical/electronic aspect, the configuration has been studied in detail to understand how the constituent components are, how they are connected to each other and what functions they have, while in the electronic control method, the general control block-schemas as well as the specific control block-schemas are presented of position, speed and current.

Electrical/Electronic Upgrades:
In the case of upgrading the system and increasing the processing capacity, the same electrical schemes will be used but with improved components such as amplifiers with higher capacity and more powerful servo motors in cases where a radical modification of the system is needed to perform the additional or heavier tasks.

Literature

1. Bilancia, P.& Schmidt, J.& Raffaeli, R.& Peruzzini, M.& Pellicciari, M. (2023). An Overview of Industrial Robots Control and Programming Approaches. Applied Sciences. 13(4):2582.

2. FANUC Robot Series. (2016). R-30iB Controller: Maintenance Manual (8th ed.). FANUC CORPORATION
3. FANUC Robot Series. (2017). R-30iB Controller: Profinet I/O Operator's Manual (4th ed.). FANUC CORPORATION
4. FANUC Robot Series. (2014). R-30iB Controller: KAREL Reference Manual (1st ed.). FANUC CORPORATION
5. FANUC Robot Series. (2016). M-900iB: Mechanical Unit Maintenance Manual (2nd ed.). FANUC CORPORATION
6. FANUC Publications. (2016). Servo Amplifier αi-B Series: Descriptions (2nd ed.). FANUC CORPORATION
7. FANUC Publications. (2015). Servo AC Servo Motor αi-B/ αi Series: Parameter Manual (9th ed.). FANUC CORPORATION
8. FANUC Publications. (2005). Servo AC Servo Motor αis/αi Series: Servo Tuning Procedure. (1st ed.). FANUC CORPORATION
9. Institute of Electrical and Electronics Engineers. Collaborative Control with Industrial Robots; Institute of Electrical and Electronics Engineers: Piscataway, NJ, USA, 2017; ISBN 9781538618837.
10. Lai, R.& Lin, W & Y. Wu, Y. (2018). Review of Research on the Key Technologies Application Fields and Development Trends of Intelligent Robots", International Conference on Intelligent Robotics and Applications, 2018.
11. Tantawi, K.H. & Sokolov, A.& and Tantawi, O. (2019). Advances in Industrial Robotics: From Industry 3.0 Automation to Industry 4.0 Collaboration. 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), Bangkok, Thailand, 2019, pp. 1-4, DOI: 10.1109/TIMES-iCON47539.2019.9024658.

The Application of Deep Learning in Optical Character Recognition

Sildi Shahini
European University of Tirana

Ardiana Topi
European University of Tirana

Forsian Elezi
University "Aleksandër Moisiu", Durrës

Abstract

Optical Character Recognition (OCR) is an essential technology for document digitization, enabling the conversion of scanned paper documents, PDFs and images into editable and searchable data. This paper focuses on the application of deep learning in OCR, particularly in digitizing handwritten medical prescriptions, where accuracy is critical for reducing errors and improving healthcare outcomes. Traditional OCR methods face challenges when dealing with handwritten texts due to the variability in handwriting styles and the quality of scanned documents. These limitations can result in recognition errors, which, in a medical context, may lead to serious consequences such as medication errors.

To address the above issue, the study explores deep learning approaches, especially Convolutional Neural Networks (CNNs), that have shown significant promise in overcoming these challenges by learning from large datasets. The study involves collecting handwritten prescriptions, preprocessing the images, and training a deep learning-based OCR model. Performance evaluation metrics, including accuracy, *precision, recall, and F1-score, indicate that the deep learning model significantly outperforms traditional OCR methods in recognizing handwritten prescriptions.*

The results demonstrate the deep learning model's ability to handle the variability of handwriting more effectively, providing a more reliable solution for digitizing medical documents. This research underlines the transformative potential of deep learning in OCR technology, particularly for critical applications such as healthcare. The findings advocate for the wider adoption of deep learning in the healthcare sector, aiming to improve patient care, reduce human error, and enhance operational efficiency, especially in pharmacy management and medical record-keeping.

Keywords: *optical character recognition, neural network, convolutional neural network, deep learning*

Introduction

Deep Learning (DL) has fundamentally changed the applications of Optical Character Recognition (OCR) technology by improving its accuracy, speed, and overall applicability. As this technology continues to evolve, integrating it with OCR will undoubtedly pave the way for new opportunities in automation and data processing efficiency across various sectors (Fateh et al., 2023).

OCR technology has undergone significant transformations since its inception, driven mainly by advancements in DL. These innovations have enabled OCR systems to achieve high accuracy in recognizing text across different document types and conditions (Nockels et al., 2022). The integration of DL into OCR gained momentum in the early 2010s when traditional methods, such as rule-based systems and basic pattern recognition, proved insufficient for complex text environments. Convolutional Neural Networks (CNNs) emerged as an especially effective model for OCR systems because of their ability to process the spatial hierarchy of images, making them ideal for handling varied layouts and font styles found in documents.

Recent studies highlight the effectiveness of DL in OCR. A standard study by Zhao et al. (2020) compared the performance of CNNs and LSTMs in OCR tasks across different languages and found that DL models consistently outperformed traditional OCR methods, with accuracy improvements of up to 20% in complex documents like invoices.

Purpose

The purpose of this thesis is to study Deep Learning technology and its application in developing a system for digitizing medical prescriptions. Medical prescriptions are critical in healthcare, containing essential information about medications, dosages, and patient instructions. However, manually processing these handwritten documents is prone to errors, time-consuming, and inefficient. This inefficiency underscores the need for an effective OCR solution that can accurately digitize handwritten prescriptions, thereby improving healthcare services, reducing medication errors, and optimizing pharmacy operations.

To address this need, the thesis uses a combination of advanced technologies, including Deep Learning, Python, Tesseract, OpenCV, and the Efficient and Accurate Scene Text (EAST) detector. By integrating these technologies, a comprehensive OCR system capable of handling the complexity of handwritten prescriptions is developed. This system combines traditional OCR techniques with modern DL approaches to achieve superior performance.

This study details the methodology for developing the OCR system, including data collection, preprocessing, model training, and evaluation. A case study on the digitization of medical prescriptions demonstrates the practical applications and benefits of this approach. By comparing the system's performance with traditional OCR methods, the thesis emphasizes significant improvements through DL technologies.

Objectives

The primary objective of this thesis is to explore and demonstrate the effectiveness of DL techniques in improving OCR for digitizing handwritten medical prescriptions. The specific objectives include:

Evaluating Traditional OCR Methods: Assessing the limitations and challenges of traditional DL techniques, especially in recognizing handwritten text in medical prescriptions.

Developing a DL-Based OCR System: Designing and implementing an OCR system using Python, Tesseract, OpenCV, and the EAST text detector to improve the accuracy and reliability of digitizing handwritten prescriptions.

Data Collection and Processing: Gathering a comprehensive dataset of handwritten prescriptions and preprocessing these images to improve text recognition.

Model Training and Optimization: Training DL models, focusing on CNNs and RNNs, to accurately recognize and digitize handwritten text.

Performance Evaluation: Evaluating the developed OCR system's performance using metrics like accuracy, precision, recall, and F1-score, and comparing it with traditional OCR methods.

Demonstrating Practical Applications: Conducting a case study on the digitization of medical prescriptions to demonstrate the practical applications and benefits of the developed OCR system.

Future Integration and Research: Exploring the potential for integrating the developed system with existing electronic health record (EHR) systems and identifying areas for future research and improvement.

Research Questions

To guide this study, the primary research question has been formulated, aiming to address the challenges and opportunities in using DL techniques to improve the accuracy and efficiency of OCR for digitizing handwritten medical prescriptions:

How can DL techniques like CNNs and RNNs be applied to improve the accuracy of OCR in recognizing handwritten text in medical prescriptions?

This question seeks to investigate the potential of advanced DL models in addressing the limitations of traditional OCR methods, particularly in processing handwritten text from medical professionals. By answering this question, the study aims to provide insights into the benefits and limitations of these techniques in the context of accurately recognizing medical text.

By answering the research questions, the thesis aims to contribute to practical and theoretical knowledge on applying DL in OCR and offer recommendations for improving healthcare services through enhanced digitization of critical medical documents.

Hypothesis

The hypothesis driving this thesis is that DL techniques, particularly those using CNNs and RNNs, can significantly improve the accuracy and efficiency of OCR for digitizing handwritten medical prescriptions compared to traditional OCR methods.

Accuracy Improvement Hypothesis: DL-based OCR will significantly improve the accuracy of recognizing handwritten medical prescriptions compared to traditional OCR methods.

The potential benefits of DL in OCR applications, especially for digitizing medical prescriptions, are substantial. The research will include comprehensive data collection, model training, and performance evaluation to determine whether DL can surpass traditional OCR methods in terms of accuracy, efficiency, and error reduction.

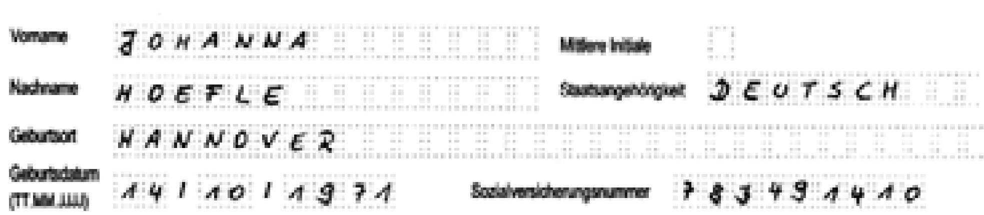
Traditional OCR Methods

Traditional Optical Character Recognition (OCR) methods have been fundamental in the early development of text recognition systems. These methods relied heavily on pattern recognition techniques, feature extraction, and template matching. Despite their initial success, traditional OCR methods encountered significant challenges, especially when dealing with handwritten text and various font styles. This section explores the key techniques used in traditional OCR and their limitations.

Pattern recognition was one of the earliest techniques employed in OCR systems. This approach involves comparing input characters with previously stored templates. Each character in the input image is matched against a database of known character patterns. Techniques like Optical Correlation were used to find the best match between the input character and stored templates. However, pattern recognition methods struggled with variations in handwriting, different fonts, and noise in scanned documents. Relying on exact or near-exact matches made these systems less robust when faced with real-world variability.

Feature extraction techniques aimed to improve the stability of OCR systems by focusing on distinguishing features of characters instead of relying on exact templates. Features like edges and angles were extracted from the input image and used to identify characters. Methods like zoning, where the character image is divided into zones and the pixel density in each zone is analyzed, became popular. The zoning method works as shown in the figure below.

FIGURE 1: Zoning Method (Cem Dilmegani, 2024)



Using statistical methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) allowed for dimensionality reduction and improved feature representation. These methods enhanced OCR systems' ability to handle variations in character appearance but still faced limitations with complex backgrounds and noise.

Template matching was another fundamental technique in traditional OCR. This approach involved comparing the input image with a set of predefined templates for each character. The character is recognized based on the template that best matches. Techniques like normalized cross-correlation were used to measure the similarity between the input image and templates. While effective for recognizing printed text with stable fonts, template matching was less successful with handwritten text due to the variability in writing styles. Moreover, template matching was computationally expensive, as it required numerous comparisons with all possible templates.

Statistical methods such as Hidden Markov Models (HMMs) and Support Vector Machines (SVMs) were introduced to improve the accuracy of OCR systems. HMMs modeled the sequential nature of text and were particularly useful for handwriting recognition. SVMs, on the other hand, provided powerful classification capabilities by finding the optimal hyperplane that separates different character classes.

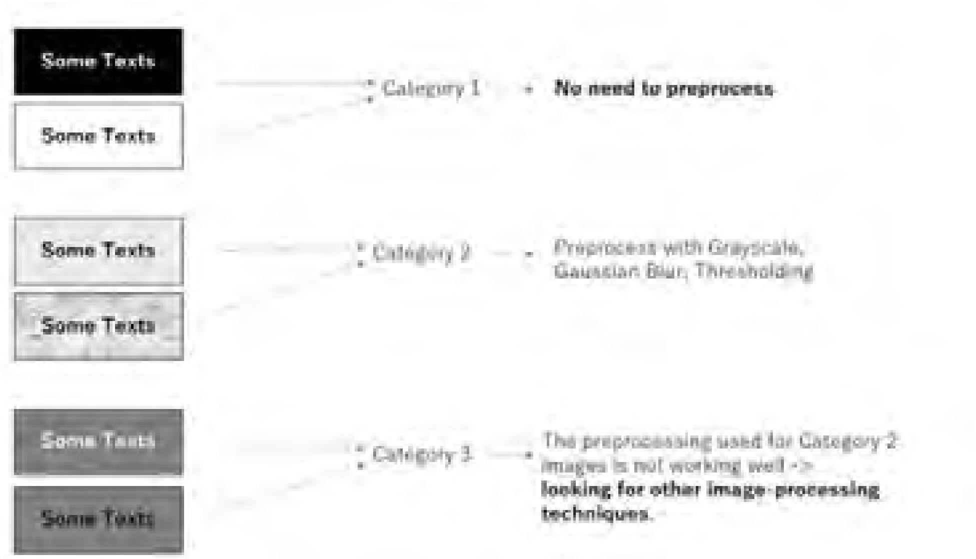
Structural methods focused on the geometric and topological properties of characters. Techniques like graph-based representations and syntactic pattern recognition exploited the structural relationships between different parts of a character. These methods aimed to offer a more holistic understanding of character shapes but required complex preprocessing steps and feature extraction.

Despite their contributions to OCR technology, traditional methods faced several limitations:

Variability in handwriting: Traditional methods struggled with the wide range of handwriting styles, leading to high error rates in recognizing handwritten text.

Sensitivity to noise: Scanned documents often contained noise, such as stains and distortions, which negatively impacted the performance of traditional OCR systems.

FIGURE 2: Noise Handling Approach (NguyenHai, 2022)



Computational complexity: Many traditional methods, particularly template matching and statistical methods, were computationally intensive, making them less suitable for real-time applications.

Limited flexibility: Traditional OCR systems were less adaptable to new fonts and handwriting styles without extensive retraining or redesign.

Traditional OCR methods laid the foundation for modern OCR systems by introducing key concepts in pattern recognition, feature extraction, and statistical modeling. However, their limitations underscored the need for more advanced techniques capable of handling the complexities of text recognition in the real world. The emergence of deep learning has addressed many of these challenges, offering more robust and accurate solutions for OCR tasks.

The Evolution of Deep Learning in OCR

The evolution of deep learning in Optical Character Recognition (OCR) has transformed the field from basic pattern recognition techniques to sophisticated and highly accurate models used today. This section reviews key moments and technological advancements that have propelled OCR forward through deep learning, highlighting recent developments from the last five years.

Early Neural Networks and OCR

The integration of neural networks into OCR began in the late 1980s and early 1990s. LeCun (1998) was one of the pioneers to introduce Convolutional Neural Networks (CNNs) for digit recognition. Their work on the LeNet-5 architecture demonstrated that CNNs could automatically extract features from images, significantly improving character recognition accuracy. This marked a departure from traditional OCR methods, which relied heavily on manually engineered features and template matching.

Convolutional Neural Networks (CNNs)

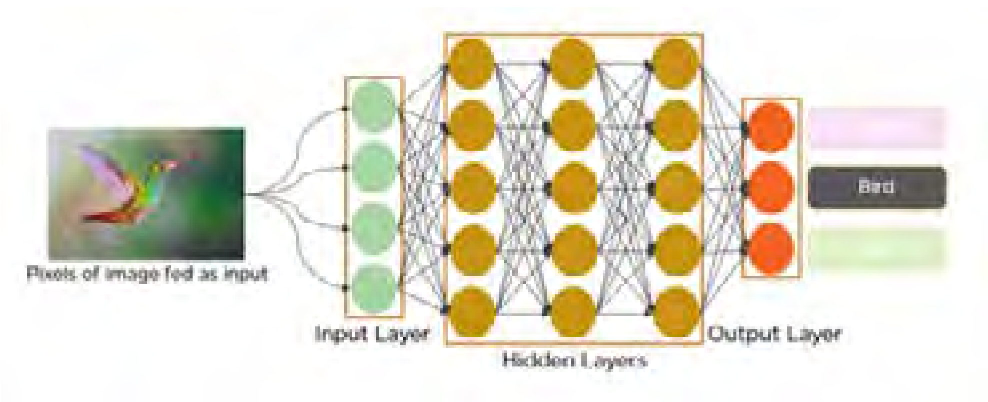
CNNs are a subcategory of machine learning at the heart of deep learning algorithms. They consist of layers of nodes, including an input layer, one or more hidden layers, and an output layer. Each node is connected to another, with an assigned weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed to the next layer.

Several types of neural networks are used for different use cases and data types. For instance, Recurrent Neural Networks (RNNs) are commonly used for natural language processing and speech recognition tasks, while Convolutional Neural Networks (CNNs) are often used for image classification and computer vision tasks. Before the advent of CNNs, manual and laborious feature extraction methods were used to identify objects in images. However, CNNs now offer a more scalable approach to image classification and object recognition tasks by leveraging principles from linear algebra, particularly matrix multiplication, to identify patterns within an image.

CNNs revolutionized OCR by providing a powerful method for feature extraction and pattern recognition. The hierarchical structure of CNNs allows them to capture spatial hierarchies in images, making them particularly effective for text recognition. Advances in deep learning frameworks like TensorFlow and PyTorch have made it easier to develop and train CNNs on large datasets.

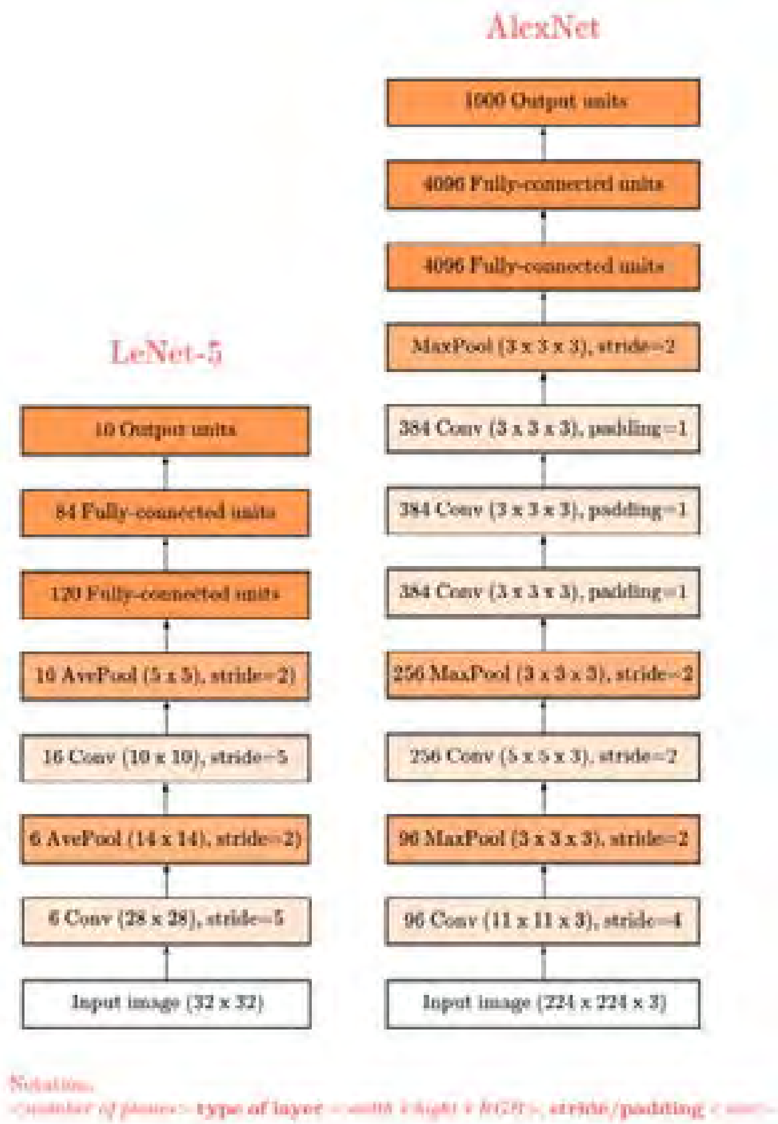
The success of the AlexNet architecture in 2012 highlighted the potential of deep learning in image processing tasks, including OCR. Developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, AlexNet was a breakthrough in deep learning. It used a deeper and wider architecture than its predecessors, with five convolutional layers and three fully connected layers. It also introduced innovations like ReLU activation functions and dropout regularization to prevent overfitting. This architecture achieved unprecedented accuracy on the ImageNet dataset, significantly outperforming traditional image recognition methods and ushering in a new era for CNNs.

FIGURE 3: AlexNet Model Performing Image Recognition Algorithm



The LeNet architecture, developed by Yann LeCun and his colleagues in the late 1980s and early 1990s, was one of the first CNN architectures designed for handwritten digit recognition. It laid the groundwork for modern CNNs with its simple but effective structure consisting of two convolutional layers followed by two subsampling layers and two fully connected layers. LeNet was instrumental in demonstrating the feasibility of using CNNs for image recognition tasks and inspired further research and development in this field.

FIGURE 4: Comparison Between LeNet and AlexNet (Pablo Caceres, 2020)



CNNs excel at handling variations in handwriting by learning from diverse datasets. Techniques like data augmentation, which involves artificially increasing the diversity of training samples, have further improved CNN robustness. This has allowed OCR systems to better generalize across different handwriting styles and text layouts.

OCR in the Field of Medicine

One of the most researched applications of OCR systems in healthcare involves the digitization of blood pressure (BP) records using Optical Character Recognition (OCR) across various forms of healthcare settings (SalahEldin Kasem et al., 2023). OCR in mobile health (m-health), particularly based on smartphone technology, is also an active research area, especially the digitization of BP records using smartphone cameras. For example, Ghoneim et al. adapted Google's TensorFlow library for mobile devices and leveraged convolutional neural networks (CNNs) in the m-health domain to digitize BP readings recorded by BP measurement instruments. This allowed for the automatic digitization of BP meter readings, ensuring a hard copy remained intact throughout the retrieval process and eliminating the need for handwritten transcription.

Another key practice in modernizing healthcare methods involves the digitization of medical charts and related materials through images or screenshots. Sokina et al. introduced a "DeepNeuro TextSpotter" (DN-TS) system that further processes bounding boxes of extracted characters and performs character segmentation, known as the textbox process. However, unlike word-spotting techniques, which require each word to be isolated and processed individually, this OCR method was more densely focused on obtaining accurate bounding boxes from medical images compared to TextSpotter.

The proposed model is reported to be evaluated using open-source projects from DeepNeuro that handle TextSpotter methods for OCR in various medical imaging languages. The evaluations include Hindi, Nepali, and Bangla, which cover most languages of the Indian subcontinent, and Arabic, English, and Malay, supporting OCR in Gulf countries and Australia.

Methodology

In this chapter, we describe the methodology followed in this study, covering the data collection process, the technology used, model training and evaluation, implementation steps, and ethical considerations involved.

The goal is to provide a clear and detailed description of the procedures and techniques used to achieve the research objectives. This chapter serves as a guide to reproduce the study and ensure that the results are reliable and valid.

The research focuses on the application of deep learning techniques in Optical Character Recognition (OCR) for the digitization of handwritten medical prescriptions. Given the critical importance of accurately extracting data in the medical field, the methodology is designed to address the unique challenges presented by handwritten text, which often includes variations in writing styles, inaccuracies, and complex medical terminology.

To achieve the research objectives, a robust and systematic approach is used, including:

Data Collection: Sourcing, anonymizing, and preprocessing a diverse set of handwritten characters and numbers.

Technologies Used: The use of advanced tools and libraries like Python, Tesseract, OpenCV, and EAST for OCR implementation.

Model Training and Evaluation: Developing and evaluating models using rigorous training processes and performance metrics.

Data Collection

We imported two datasets for the training phase, one for numbers (MNIST 0-9 from Keras) and one for letters (Kaggle). These datasets consist of handwritten numbers and letters, with over 70,000 images of numbers and 372,550 letter images. The datasets were imported and installed as follows:

For numbers: Keras – MNIST 0-9

For letters: Kaggle A-Z

Data Preprocessing

The raw data collected from healthcare institutions goes through several preprocessing and annotation steps to prepare it for model training. These steps are crucial for improving the image quality and creating a true dataset for accurate model evaluation:

- *Digitization and Scanning*: All handwritten characters are scanned using high-resolution scanners to convert them into digital images, ensuring clarity and suitability for OCR processing.
- *Image Preprocessing*: Scanned images undergo various preprocessing techniques to enhance their quality and readability, such as:
  - *Binarization*: Converting images to a binary (black and white) format to improve text visibility.
  - *Noise Reduction*: Removing noise and background artifacts to ensure clean text areas.
  - *Image Normalization*: Standardizing image sizes and resolutions to ensure uniformity across the dataset.
- *Manual Annotation*: Each character is manually labeled to create a true dataset, including:

    - *Text Labeling*: Manually transcribing handwritten text into digital format, ensuring accuracy and consistency.
    - *Zone Marking*: Identifying and marking areas of interest within the images.
- *Quality Assurance*: A quality assurance process is implemented to verify the accuracy of the annotations.

Data Augmentation

To increase dataset diversity and improve the model's ability to generalize, data augmentation techniques are applied. These techniques simulate variations that the model might encounter in real-world scenarios. Augmentation methods include:

- *Rotation*: Rotating images to simulate different writing angles.
- *Resizing*: Adjusting image sizes to account for variations in character formats.
- *Image Cropping*: Randomly cropping parts of images to imitate occlusions and partial text.
- *Lighting and Contrast*: Modifying lighting and contrast to reflect different scanning conditions.

Technologies Used

This section details the technologies utilized in the project to digitize handwritten medical prescriptions using **Optical Character Recognition (OCR)**. The chosen technologies provide a powerful, efficient, and accurate system for OCR, combining advanced image processing and deep learning tools.

- *Python*: Python was the primary programming language due to its simplicity, versatility, and extensive ecosystem of libraries for machine learning and image processing. It facilitates fast development and testing, with access to powerful libraries like **TensorFlow, Keras, OpenCV, and Tesseract**.
- *Tesseract*: Tesseract is an open-source OCR engine known for its accuracy in recognizing both printed and handwritten text. Initially developed by HP and now maintained by Google, it supports 166 languages, including Albanian. It integrates easily with Python using pytesseract and allows for customizations to improve recognition accuracy.
- *OpenCV*: OpenCV is a widely used open-source computer vision library for image processing tasks. In this study, it aids in preprocessing images for OCR, handling tasks like **binarization**, **noise reduction**, and **normalization**.

Combined with Tesseract, OpenCV enhances text recognition from scanned images.
- *EAST (Efficient and Accurate Scene Text Detector)*: EAST is a deep learning-based text detector designed for identifying text in complex scenes. It is highly accurate, fast, and suitable for real-time applications. In this project, EAST helps detect text areas in images, allowing Tesseract to accurately recognize and convert them into structured digital formats.

Training and Evaluation

Training and evaluation are critical to ensuring the success of the OCR system in recognizing handwritten medical prescriptions. This section outlines the detailed processes of preparing the dataset, training the model, evaluating performance, and validating the results to ensure accuracy and reliability.

The training process involves several key steps to ensure the models learn effectively from the data and generalize well. The main focus is on optimizing the **EAST text detector** and **Tesseract OCR engine** using a **Convolutional Neural Network (CNN)** architecture.

- *Dataset Preparation*:
  - **Training Set**: Consists of 2,700 images, combining digits (0-9 from MNIST) and English alphabets (a-z from Kaggle). These images are repeated 20 times to enhance training coverage.
  - **Validation Set**: A subset used to adjust hyperparameters and prevent overfitting.
  - **Test Set**: A separate dataset reserved for final evaluation, ensuring unbiased model assessment.

Model Training:

- *EAST Text Detector*:
  - Uses a CNN architecture designed to detect text at the pixel level.
  - **Loss Function**: Combines classification loss (text/non-text regions) and regression loss (bounding box coordinates).
  - **Hyperparameters**: Includes learning rate, batch size, and epochs, optimized through grid search and cross-validation.
- *CNN Layers*:
  - Involves convolutional layers for feature extraction, pooling layers for down-sampling, and fully connected layers for classification.
  - **Activation Functions**: Uses ReLU for non-linearity and softmax for output classification.

    - **Optimizer**: Adam is used to fine-tune learning rate and improve convergence.
- *Tesseract OCR*:
    - **Pre-trained** on a large corpus of printed and handwritten text.
    - Further trained on specific datasets to adapt to different handwriting styles.
    - Custom configurations are applied for segmentation and recognition of variable writing styles.

Several evaluation metrics are used to assess the performance of OCR models:

- *Accuracy*: Measures the proportion of correctly recognized text instances out of the total.
- *Precision*: Indicates the proportion of correctly recognized text out of all recognized instances, reflecting the model's ability to avoid false positives.
- *Recall*: Measures the model's ability to detect all true text instances.
- *F1-Score*: Harmonic means of precision and recall, balancing both aspects.
- *Character Error Rate (CER)*: The ratio of incorrect characters to the total number of characters, reflecting detailed recognition accuracy.
- *Word Error Rate (WER)*: The ratio of incorrect words to the total number of words, providing overall text recognition accuracy.

Methods and analysis

Preprocessing Scripts

Preprocessing is an essential step in the image analysis and computer vision tasks pipeline. It involves transforming raw image data into a format more suitable for further processing and analysis. This section covers the basic steps required to preprocess images using OpenCV, a well-known library for computer vision applications. The following steps are critical for effective preprocessing:

- *Loading Images***:** Images are loaded into the program from the file system.
- *Converting to Grayscale***:** The loaded images are converted to grayscale to reduce complexity and focus on intensity information.
- *Thresholding and Binarization***:** Thresholding techniques are applied to convert grayscale images into binary images, making the object of interest and background easily distinguishable.
- *Noise Reduction***:** Noise is removed from the binary images to improve the accuracy of further processing steps.

- *Normalization*: Pixel values are normalized to a standard scale to ensure consistency across different images.

Thresholding

Thresholding is an image segmentation method. For example, in an image containing both a dog and a tree, we need to extract the dog. This is done by segmenting the image into areas of interest and non-interest using a point called the threshold.

- **0 - Black:** RGB (0, 0, 0) - black
- **255 - White:** RGB (255, 255, 255) - white

When applying thresholding, we filter the pixels. For instance, with a threshold of 127, pixels with values >= 127 will turn white, while those with values less than 127 will turn black (0). This technique is called binarization, meaning we only have two values: 0 for black and 1 for white. Binarization simplifies the image, making it easier for the algorithm to detect objects, such as text, by reducing complexity.

Gaussian: Uses the mean along with the standard deviation, applying convolution. This technique produces better results by smoothing out noise and making thresholding more adaptive to variations in the image. Gaussian thresholding dynamically adjusts the threshold based on the local region, improving accuracy, especially in images with uneven lighting or noise.

We have chosen **Gaussian Thresholding** because it provides stable and smooth results in noisy images by using the mean and standard deviation for better processing. This method adjusts the threshold dynamically across different regions of the image, making it particularly effective in handling variations in lighting or noise, resulting in more refined segmentation.

Noise Reduction

Morphological operations are techniques used for noise removal, edge detection, and image enhancement. These methods are applied exclusively to binary images (black and white), where 0 represents black and 1 represents white.

Main Operations:

- *Erosion and Dilation*:
  - **Erosion:** Removes pixels, making the image thinner (e.g., thinning text).
  - **Dilation:** Adds pixels, making the image bolder (e.g., emphasizing text).

- *Opening*:
  - This involves applying erosion followed by dilation. First, pixels are removed, then new pixels are added.
  - **Usage:** Mainly used for noise removal. During erosion, unwanted pixels are removed, thinning the text. Dilation then restores the text to its original shape, enlarging the edges.
- *Closing*:
  - This involves applying dilation followed by erosion. First, new pixels are added, and then some are removed.
  - **Usage:** Useful for removing noise within the image (e.g., colored spots inside a letter). When dilation is applied, these spots disappear, and the image edges are enlarged. Erosion is then applied to remove excess pixels.

Text Detection Using EAST

In this section, we will explore the process of detecting text within images using the **Efficient and Accurate Scene Text (EAST)** detector. This method is particularly effective for detecting text in complex scenes, as it leverages a deep learning model to accurately identify text areas. The EAST model provides a robust solution for real-time applications due to its efficiency and precision.

The first step in using the EAST detector is loading the pre-trained model. The EAST model is available as a pre-trained neural network in OpenCV, making it easy to integrate into your text detection pipeline.

Once the model is loaded, the next step is processing the images to detect text regions. This involves creating a **blob** from the image, inputting it into the network, and feeding it forward through the network to retrieve the text regions. The images should be resized to dimensions of **320 x 320 (width x height)** for optimal detection. If the images are of a different size, resizing must be performed to ensure consistent input dimensions.

FIGURE 13: Region of interest detection

Pre-processing

Processing

Text Recognition with Tesseract

In this section, we will examine the process of recognizing text from images using **Tesseract**, an open-source OCR (Optical Character Recognition) engine known for its accuracy and flexibility. After detecting text regions with EAST, the next step is extracting and recognizing the textual content within these regions.

Tesseract allows us to configure various parameters to meet the specific needs of our dataset, ensuring more accurate text recognition. This customization is key to achieving high performance, especially when working with different fonts, languages, or noisy images.

FIGURE 15: Detektimi i tekstit

Training the OCR Model

This section describes the process of training an Optical Character Recognition (OCR) model. The training was performed using a **Convolutional Neural Network (CNN)** model and two well-known datasets: **MNIST** (containing digits from 0 to 9) and **Kaggle A-Z** (containing letters from A to Z). A total of **2700 images** from these datasets were used for training, with each image being augmented 20 times to expand the training dataset and improve results.

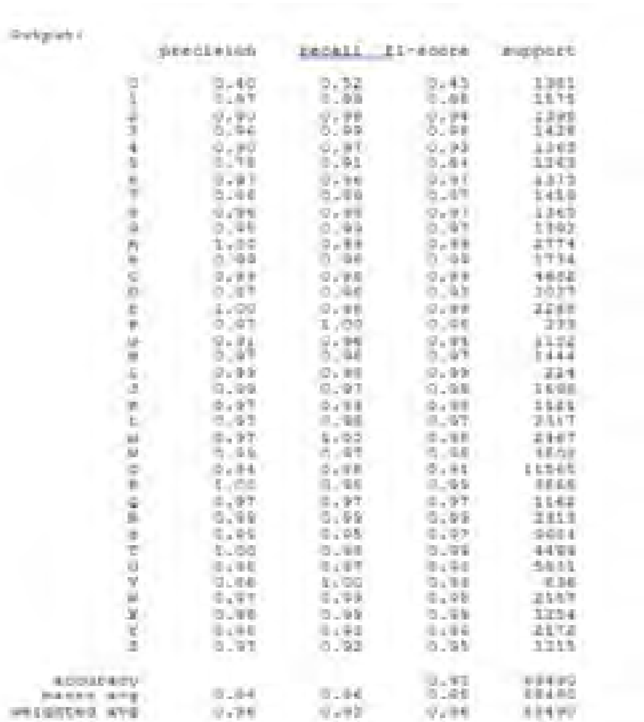
This section will detail the following steps:

- **Dataset Preparation:** Organizing and preprocessing the MNIST and Kaggle A-Z datasets for training.
- **CNN Model Configuration:** Setting up the CNN architecture to effectively recognize characters.
- **Training Process:** Feeding the model with the training data and running it through multiple epochs.
- **Validation Methods:** Ensuring the model's performance is tested and validated to generalize well to new data.

Through this process, the goal is to achieve an OCR model capable of recognizing characters from the given datasets with high accuracy and reliability.

Trajnimi i rrjetit neural

FIGURE 16: Error Graph



	precision	recall	f1-score	support
accuracy				
macro avg				
weighted avg				

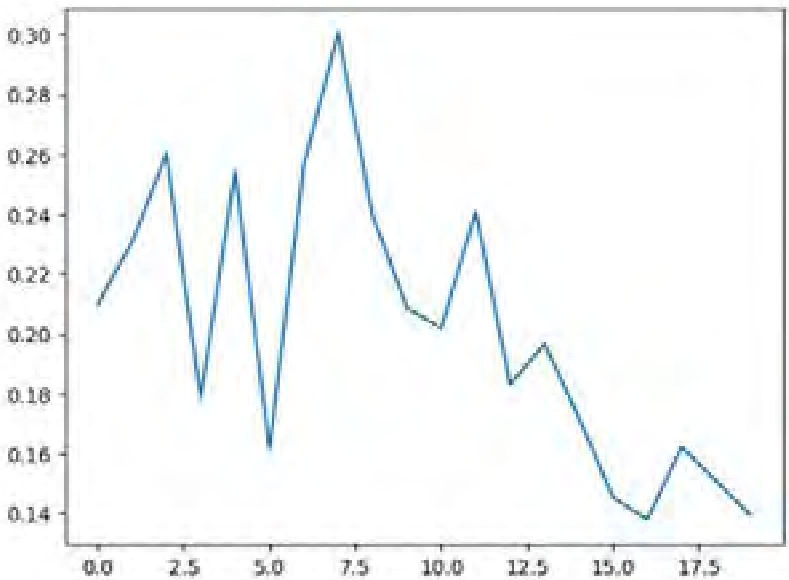
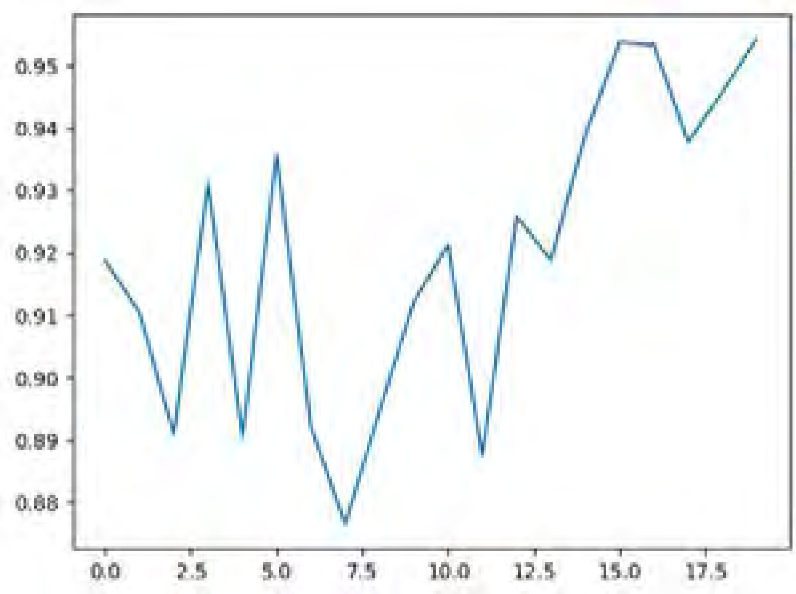


FIGURE 17: Graph of Accuracy



Results

In this chapter, we will explain the application of the trained OCR model to digitize handwritten medical prescriptions from doctors. By utilizing the trained model, we will test its performance on an actual doctor's prescription to evaluate its accuracy and effectiveness in real-world scenarios. This testing will help assess how well the model can interpret and convert handwritten text into digital format, particularly focusing on the unique challenges posed by messy handwriting, abbreviations, and medical terminology typically found in prescriptions.

FIGURE 18: Test image

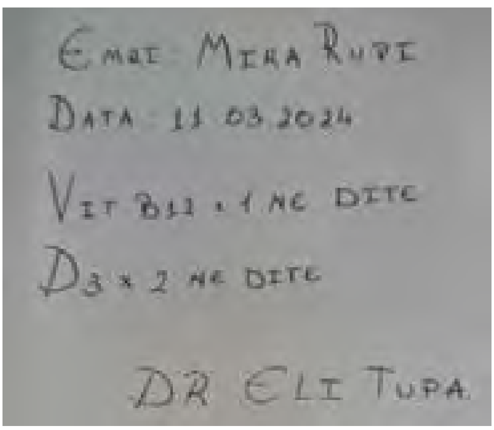
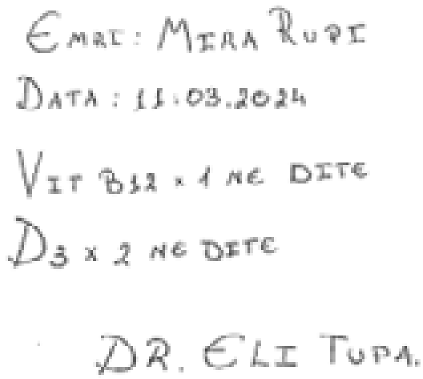
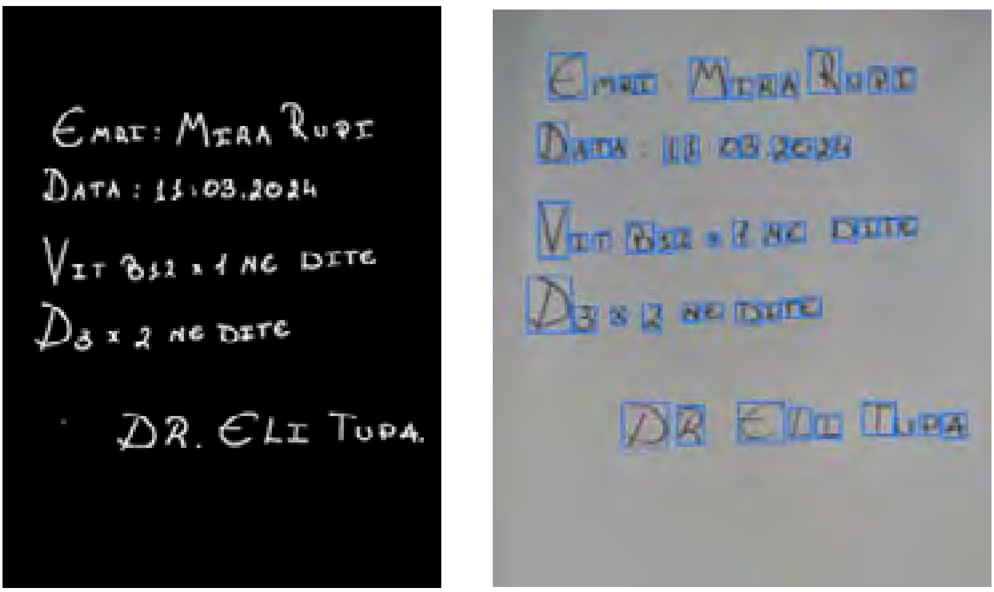


FIGURE 19: Preprocessing of the image

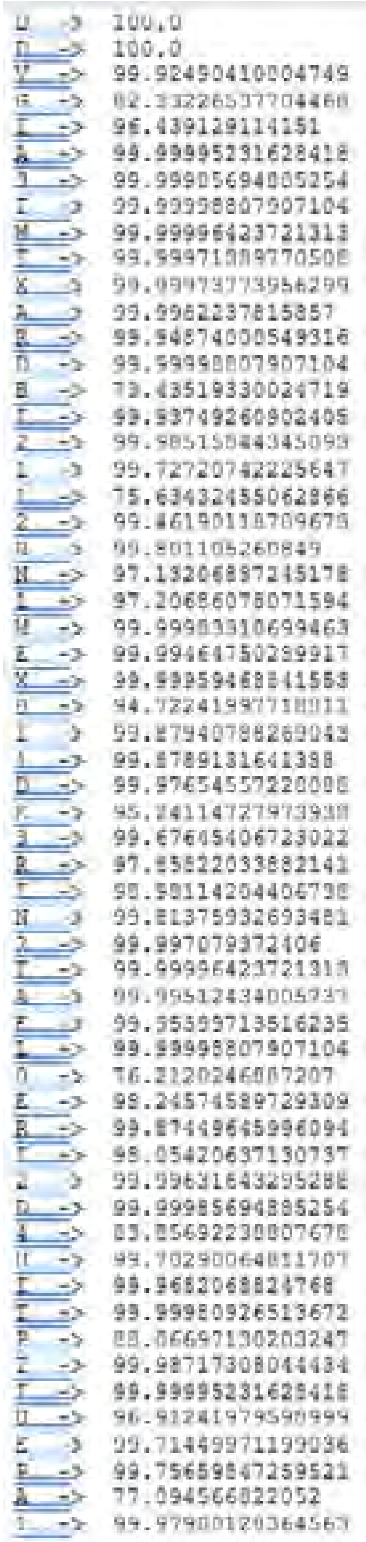


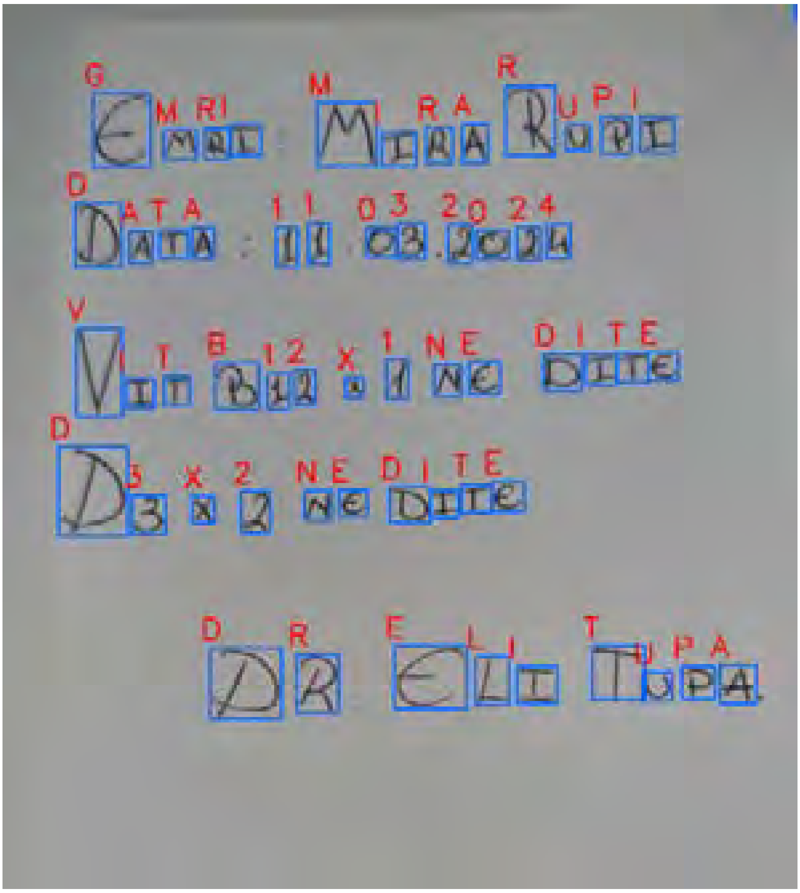
We have separated the text from the background, then we invert the colors so the background is black and text in white because that's how the neural network is trained. After we apply Dialation to make the text more highlighted.

FIGURE 21: Confidence



Letter detection





Conclusion

This study explores the application of Deep Learning technologies, with a focus on Convolutional Neural Networks (CNNs), to enhance the accuracy and efficiency of Optical Character Recognition (OCR) for handwritten medical prescriptions. The research demonstrates that combining CNNs, which extract complex features from images, with Recurrent Neural Networks (RNNs), which manage long-term dependencies in character sequences, significantly improves OCR performance. By training models on diverse datasets and using data augmentation techniques, the study achieved a CNN model accuracy of 95.28%, with only one character error, marking a substantial improvement over traditional OCR method.

The results indicate that deep learning models not only handle variations in handwriting more effectively but also automate the OCR process faster and more reliably. This research confirms the transformative potential of deep learning technologies for digitizing handwritten documents, especially in the medical field.

Furthermore, the study suggests that such technologies can be applied to improve the accuracy and efficiency of OCR in other areas beyond medical prescriptions, offering widespread practical applications.

Bibliography

1. Caceres, P. (2020). The Convolutional Neural Network - Theory and Implementation of LeNet-5 and AlexNet. Available at: pabloinsente.github.io
2. Dilmegani, C. (2024). State of OCR in 2024: Is it dead or a solved problem?
Available at: https://research.aimultiple.com/ocr-technology/
3. Fateh, A., Fateh, M., & Abolghasemi, V. (2023). Enhancing optical character recognition: Efficient techniques for document layout analysis and text line detection. Engineering Reports. https://doi.org/10.1002/eng2.12832
4. Hai, N. (2022). How to verify if the image contains noise in background before 'OCR' ing.
Available at:
https://stackoverflow.com/questions/70604177/how-to-verify-if-the-image-contains-noise-in-background-before-ocr-ing
5. Kasem, S.E. (2023). Advancements and Challenges in Arabic Optical Character Recognition: A Comprehensive Survey. Available at: https://arxiv.org/pdf/2312.11812
6. Nockels, J., Gooding, P., Ames, S., & Terras, M. (2022). Understanding the application of handwritten text recognition technology in heritage contexts: a systematic review of Transkribus in published research. Archival Science, 22(3), 367–392. https://doi.org/10.1007/s10502-022-09397-0.
7. Zhao, Z., Jiang, M., Guo, S., Wang, Z., Chao, F., & Tan, K. C. (2020). Improving Deep Learning based Optical Character Recognition via Neural Architecture Search. In 2020 IEEE Congress on Evolutionary Computation, CEC 2020 - Conference Proceedings (pp. 1-7). Article 9185798 (2020 IEEE Congress on Evolutionary Computation, CEC 2020 - Conference Proceedings). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CEC48606.2020.9185798