EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"

Research in Computing Science

Vol. 152 No. 1
January 2023

Research in Computing Science

Research in Computing Science, Año 22, Volumen 152, No. 1, enero de 2023, es una publicación mensual, editada por el Instituto Politécnico Nacional, a través del Centro de Investigación en Computación. Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, Ciudad de México, Tel. 57 29 60 00, ext. 56571. https://www.rcs.cic.ipn.mx. Editor responsable: Dr. Grigori Sidorov. Reserva de Derechos al Uso Exclusivo del Título No. 04-2019-082310242100-203. ISSN: en trámite, ambos otorgados por el Instituto Politécnico Nacional de Derecho de Autor. Responsable de la última actualización de este número: el Centro de Investigación en Computación, Dr. Grigori Sidorov, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738. Fecha de última modificación 01 de enero de 2023.

Research in Computing Science, year 22, Volume 152, No. 1, January 2023, is published monthly by the Center for Computing Research of IPN.

Geographic Information Systems Latin America (GISLATAM)

Roberto Eswart Zagal Flores
José Antonio León Borges
Miguel Félix Mata Rivera (eds.)

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2023

ISSN: in process

Formerly ISSNs: 1870-4069, 1665-9899

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional "Adolfo López Mateos", Zacatenco
07738, México D.F., México

http://www.rcs.cic.ipn.mx
http://www.ipn.mx
http://www.cic.ipn.mx

Indexed in LATINDEX, DBLP and Periodica

Electronic edition

ISSN 1870-4069

Table of Contents

Page

WEB-GIS Platform to Visualize Small and Medium Enterprises with Circular Business Models in Mexico 5
Diana G. Castro-Frontana

Hazard Identification Procedure, Risk Analysis, and Determination of Controls for the Failure of a Line Valve in a Gas Pipeline 23
Vladimir Avalos-Bravo, Adriana Guzmán-Ruvalcaba, Patricia Alejandra Vallejo-Landa

The application of Machine Learning in the Develop of Mobile Applications for Geographic Information Systems Using Create ML, CoreML and Mapkit for iOS 39
Eduardo Eloy Loza Pacheco, Mayra Lorena Díaz Sosa, Christian Carlos Delgado Elizondo, Mayra Olguin Rosas

Implementation of an Embedded System to Detect the Use of Face Masks Using a CNN Model for the Networking Laboratory at the UQRoo Campus Cancun 47
María Kristell Flores Reyes, Daniel Sergio Martínez-Ramírez, Julio Cesar Ramírez-Pacheco, Homero Toral-Cruz, David Ernesto Troncoso-Romero, Ismael Osuna-Galán, J. A. León-Borges

ISSN 1870-4069

WEB-GIS Platform to Visualize Small and Medium Enterprises with Circular Business Models in Mexico

Diana G. Castro-Frontana

Instituto Politécnico Nacional,
Escuela Nacional de Ciencias Biológicas,
Mexico

dgcastro@ipn.mx

Abstract. The Circular Economy is a visionary concept for a new economic paradigm. It is founded on three fundamental principles: I) To regenerate nature by shifting our economy from linear to circular to preserve and enhance natural capital by controlling finite reserves and balancing renewable resource flows. II) To optimize resource yields by distributing products, components, and materials with their maximum utility in both technical and biological cycles (to pre-serve an extended life of products; and responsible use and manufacturing). III) To promote system effectiveness by detecting and eliminating negative externalities from the design (use waste as a resource). The Circular Economy advocates for specific measures to advance this model, encompassing additional actions beyond the established 3R principle of "Reduce, Reuse, and Recycle". These additional actions are known as the R-ladder or the 7-R hierarchy. This hierarchy was used to filter economic units from the 2019 Economic Census in Mexico (released by INEGI) to classify micro, small and medium-sized enterprises (MSMEs) that perform activities related to the circular economy principles. In fact, MSMEs companies (MIPyMES for their acronym in Spanish) dominate the nation's commercial landscape, comprising 99.8% of all businesses. The intention of this research is to create a web-GIS platform to visualize and empower businesses, especially micro, small and medium companies, that are contributing and will keep contributing to achieve the three circular economy principles. A web-GIS platform prototype was created using ArcGIS Pro soft-ware and the ArcGIS Online platform. According to the results, a total of 44,943 micro, small, and medium-sized companies across three Mexican states are involved in Circular Economy practices, although many remain unaware of their participation in this model. These companies can foster improved, sustainable employment opportunities. Efforts are underway to process and incorporate data from an additional 29 states into the Circular Economy Platform. Similar procedures will be implemented for the remaining Mexican states to ensure comprehensive nationwide coverage.

Keywords: Micro, small, medium companies, MSMEs, MIPyMES, economic units, circular economy, WEB-GIS platform.

1 Introduction

1.1 Circular Economy

Our existing economic model adheres to linear production patterns, characterized by production, consumption, and waste, fostering an unsustainable environment that ac-celerates the depletion and degradation of our natural resources. It is within this context that the proposal for a Circular Economy model emerges. The Circular Economy is defined as "*an economic system designed to uphold the continued utility and value of products and their components, fostering a continuous cycle of sustainable develop-ment that safeguards and enhances natural capital, while optimizing the use of finite resources*"[1]. The Circular Economy (CE) is founded on three principles:

Principle 1: To regenerate nature by shifting our economy from linear to circular to preserve and enhance natural capital by controlling finite reserves and balancing renew-able resource flows.

Principle 2: To optimize resource yields by distributing products, components and materials with their maximum utility at all times in both technical and biological cycles (preserve an extend life of products; and responsible use and manufacturing of products).

Principle 3: To promote system effectiveness by detecting and eliminating negative externalities from the design (use waste as a resource).

Fig. 1 shows in detail what circular economy involves. It is known as the butterfly diagram.

According to the above principles, one of the main guidelines of the circular economy is to ensure that products, materials, and resources have a longer life and are not destroyed but reused. With this, it will be possible to eliminate a large percentage of waste that ends up in open dumps and to eliminate CO_2 emissions. By promoting the reuse of materials and energy, harmful impacts on the health of current and future generations are reduced and natural capital is preserved. Through a circular approach then, CE can significantly contribute to reducing waste, increasing resource efficiency, and promoting sustainable development [2, 3].

To successfully implement the circular economy model, it is imperative to engage various sectors of production, ranging from policymakers and legislators to end consumers who procure and dispose of finished products.

This transition requires a concerted effort toward a systemic and comprehensive transformation. As illustrated in Fig. 1, the Circular Economy advocates for specific measures to advance this model, encompassing additional actions beyond the established 3R principle of “Reduce, Re-use, and Recycle”.

The rise of ecological concern has made the 3R principle evolve to currently 7Rs (Reduce, Reuse, Recycle, Redesign, Renew, Repair and Recover), also known as the 7-R hierarchy [4]. Another version of these additional actions can be organized in what is known as the R-ladder [3]. It can be used as a checklist for designers and innovators (see Fig 2). This ladder is used to prioritize strategies and actions towards

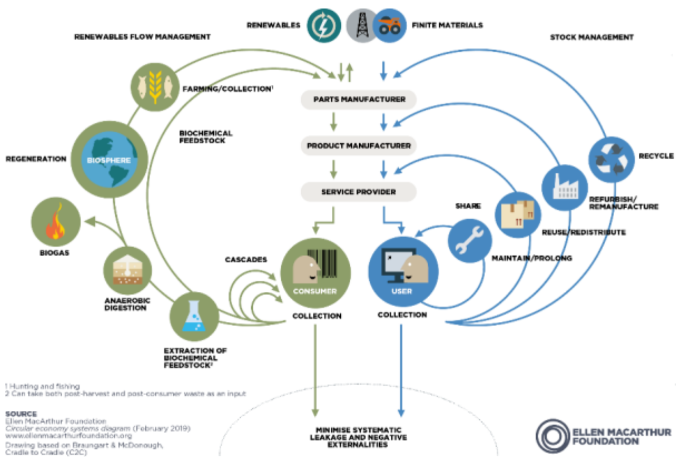

Fig. 1. The butterfly diagram explains in detail the three principles of circular economy.

Fig. 2. The R-Ladder [3] also known as the circular economy hierarchy.

circular economy so that the higher R up on the ladder, the more impact a solution has.

1.2 Directory of Economic Units (DENUE)

The National Statistical Directory of Economic Units (DENUE) is the minimum infrastructure of the National Economic Information Subsystem (SNIE) whose purpose is

Fig. 3. Number of establishments and enterprises censed by INEGI in 2023.

Table 1. Businesses classification by the number of employees.

Size	Number of employees	Annual sells (range) in Mexican pesos
Micro	1 - 10	4 to 4.6 million MXN
Small	11 - 30	40 to 100 million MXN
Medium	31 - 100 (commercial) 51 - 100 (services)	100 to 250 million MXN

to provide both, specialized and non-specialized users, with the identification, location, and contact data of the economic units active in the national territory for decision-making support and to optimize resources in both the public and private spheres. This data was also created to facilitate the development and assessment of public policies and economic development programs at all three tiers of government. The DENUE was created, and it is managed and updated by the Mexican National Institute of Statistics and Geography (INEGI). There, economic units are cataloged and categorized into two distinct groups: establishments and enterprises [5].

- **Establishment:** this category refers to an economic unit firmly rooted in a solitary physical location, demarcated by fixed structures and installations, and combining efforts and resources under the governance of a single owner or controlling entity to conduct various economic activities, whether for profit or not. This category includes domiciles where economic activities are undertaken, while excluding those pursued solely for self-consumption, and services rendered off-site. Also, information regarding activities in the manufacturing industry, commerce, private non-financial services, and financial and insurance services pertains to this unit.
- **Enterprise:** This refers to an organization, owned by a single legal entity, engaged in one or more economic activities, possessing autonomy in marketing,

financing, and investment decisions, and vested with the authority and responsibility to allocate resources in accordance with a production plan or strategy for goods and services. Such entities may be situated or operate across multiple locations.

The Directory excludes economic units conducting itinerant activities or those with demountable premises that are dismantled daily. One of the major advantages of using the DENUE database is that the directory is continuously updated by authorized in-formants, who can update or supplement the data of their businesses and incorporate commercial information online through the application found in the technical file of each economic unit. This information was previously validated by INEGI [6].

The economic units are already georeferenced. INEGI offers the information for free at its website and it can be downloaded in cvs or shape file format. This way, it was easier to manage data by state. The North American Industry Classification System (NAICS) is the standard used by Federal statistical agencies in classifying business establishments for the purpose of collecting, analyzing, and publishing statistical data related to the U.S [7]. It has also become the standard for Mexico and Canada because of the free trade agreement (NAFTA) among these three countries. So, it is mandatory for INEGI to publish statistical data about economic activities using this standard [8].

1.3 Micro, Small and Medium Business (MSMEs)

According to INEGI [9], micro, small and medium-sized businesses are given their category by the number of their employees, as shown in table 1.

Based on the 2019 Economic Census results released by INEGI, micro, small and medium-sized enterprises (MSMEs) dominate the nation's commercial landscape, comprising 99.8% of all businesses. Among the 4,773,995 economic units scrutinized across both the private and parastatal sectors by INEGI, 95% fall within the micro-sized category (0 to 10 employed persons), 4% fall within the small-sized category (11 to 50 persons), and 0.8% are classified as medium-sized enterprises (51 to 250 persons). Merely 0.2% are categorized as large-scale enterprises (exceeding 251 persons) [9].

Moreover, MSMEs collectively contribute 68.4% to the overall tally of businesses, with micro-businesses accounting for 14.2% of the total generated income, small businesses for 16.1%, and medium-sized businesses for 21.9%. Consequently, MSMEs are responsible for generating 52.2% of the total income derived from commercial activities.

1.4 Sustainable Development Goals (SDG)

In 2015, during the Sustainable Development Summit held from September 25 to 27 in New York, the countries part of the UN approved the 2030 Agenda in which the 17 Sustainable Development Goals (SDGs) are established (see Fig.4). The 2030

Fig. 4. Sustainable Development Goals [10].

Agenda sets out these 17 Goals with 169 integrated targets covering the economic, social and environmental spheres. It will govern global development programs for the next years.

The States signed up to it are committed to mobilizing the necessary resources for its implementation through partnerships focused especially on the needs of the poorest and most vulnerable [6].

The Circular Economy is not indifferent to the Sustainable Development Goals (SDGs) or the 2030 Agenda; on the contrary, it can help to achieve it by supporting job creation, counteracting social problems arising from unemployment, and promoting social sustainability. Circular economy also has a strong potential to transform society toward environmental, economic, and social sustainability [11]. This project looks for contributing to the following SDG by generating information to these specific targets:

- SDG 8: Promote sustained, inclusive, and sustainable economic growth, full and productive employment, and decent work for all (targets 8.2, 8.3, and 8.4).
- SDG 9: Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation (targets 9.2, 9.3, 9.5, 9.b and 9n.3).
- SDG 11: Make cities and human settlements inclusive, safe, resilient and sustainable (target 11.6).
- SDG 12: Ensure sustainable consumption and production patterns (target 12.5 and 12n.1).

Table 2. Some SDG that this project seeks to contribute to, specific targets and their indicators.

SDG	Target	Indicator
SDG 8	8.4 Progressively improve, by 2030, the efficient production and consumption of global re-sources and strive to decouple economic growth from environmental degradation, in line with the 10-Year Framework of Programs on Sustainable Consumption and Production Patterns, starting with developed countries.	8.4.2.a Domestic material consumption in absolute terms.
SDG 9	9.n3 Increase the contribution of micro, small, and medium-sized enterprises to employment and GDP.	9.n3.1 Percentage of employed personnel generated by MSMEs (includes salaried, non-salaried and jobs provided by another co-pany).
SDG 11	11.6. By 2030, reduce the per capita negative environmental impact of cities, paying special attention to air quality and municipal and other waste management.	11.6.1.a Percentage of municipal solid waste collected and disposed of appropriately.
SDG 12	12.5 By 2030, significantly reduce waste generation through prevention, reduction, recycling and reuse activities. 12n.1 Promote the circular economy in production and consumption chains, understood as the redesign of products and services to reduce waste at the end of their useful life and from a shared value perspective.	12.5.1.a Percentage of municipal solid waste collected that can be recycled. 12n.1.1 Percentage of municipalities with adequate disposal of urban solid waste

– SDG 13: Take urgent action to combat climate change and its impacts (target 13.3).

An example of how this platform can help create information to achieve SDG at local level is shown Table 2. INEGI is also responsible for monitoring SDG in Mexico.

All the targets and specific indicators and their evolution through years is of public access at the web site Agenda 2030 [11].

2 Research Objectives

To develop a geographically referenced GIS/Web platform prototype aimed at visualizing and advancing businesses dedicated to the reuse, remanufacturing, recycling, etc., of products, from DENUE datasets and enriched by crowd mapping.

2.1 Specific Objectives

- To classify and refine the DENUE platform to extract data concerning micro, small, and medium-sized enterprises actively involved in any of the R-ladder activities (product repair, remanufacturing, and recycling, etc.).
- To map out the businesses identified in the analysis of the DENUE on an online GIS platform to visualize the micro, small and medium-sized companies involved in the circular economy ladder activities (R ladder) in Mexico.
- To establish the pioneering platform in Mexico that advocates and consolidates business information encompassing all facets of the Circular Economy (R hierarchy).

3 Materials and Method

The materials used for this project are:

- Database of the National Statistical Directory of Economic Units (DENUE). Version 2019 (datasets) in vector files (georeferenced points).
- Map Digital Software version 6.3. Open software (v6.3) developed by INEGI.
- ArcGIS Pro Software and ArcGIS Online (for the creation of the WEB platform)

For the development of this research, the following steps were involved:

1. The DENUE database was downloaded from the INEGI website, for entities of the Mexican Republic.
2. Only three entities were selected for the web-platform prototype: Mexico City, Hidalgo and Aguascalientes. There was no specific reason to choose these states except that we wanted to include at least a state with large (Mexico City), medium (Hidalgo) and small number of economic units (Aguascalientes).
3. All the original data in DENUE is classified according to the NAICS standard used in the US, Mexico and Canada. Codes from the North American Industry Classification System (NAICS) were analyzed to find potential activities related to Circular Economy. For example, shoes repairing has the NAICS code: 316214. Repair of leather footwear and other leather goods has the NAICS code: 811430.
4. Then, these codes were filtered in the DENUE database of each state to save them in another file containing only these registers. Some fields from the dataset with no relevant information were eliminated. The databases were purged, eliminating all but the following categories:
   - Rental,
   - Alignment and balancing,
   - Confection,
   - Industrial design,

- Waste management,
- Repairs,
- Cleaning services,
- Tire revitalization (vulcanizers),
- Remediation,
- Treatment and final disposal.

5. A new field for each dataset was created. The new field contained a new R-category assigned to each company (depending on its original NAICS code and its activities and the R in the R-ladder that best suits it).
6. Additionally, some new categories such as public infrastructure for wastewater treatment, remediation, collection, and hazardous waste treatment were created to sub-classify the database. This, to differentiate activities that are not part of the R-ladder but could contribute to circular economy goals.
7. The "Rental" category was created (for all business codes regarding rental activities).
8. The "Collection and/or recycling" category was created to group solid waste companies that perform one or both activities.
9. The database was reviewed again, and certain companies (company name) were randomly selected to verify on the web what they do and to refine their classification in one of the created categories.
10. The verification was performed to check if certain businesses still existed. Google Streetview was used to corroborate that the business exists (because in the pandemic some businesses may have closed). It is important to mention that this verification was done during the Covid-19 pandemic lockdown period in 2021.
11. Once the records were reclassified into one of the R categories, the files were ex-ported as new georeferenced files (shapefiles) to integrate them into a GIS online platform.

3.1 Integration of Layers to the Platform

12. The suite ArcGISPro (which includes ArcGIS Online) was used to create the web platform prototype. The reclassified database (vector file) was uploaded as a point layer (each point represents an economic unit) and each layer contains all the data for a particular state (Mexico City, Aguascalientes or Hidalgo).
13. Each layer was added a legend to identify the R-ladder categories with a single color each.

Table 3. Economic units (MSMEs) before and after the reclassification process.

Mexican State	**Economic Units in**	**Economic units (MSMEs) after**
	Denue 2019	**Reclassification using the R-ladder**
Mexico City	476553	30607
Hidalgo	120570	10175
Aguascalientes	57840	4161

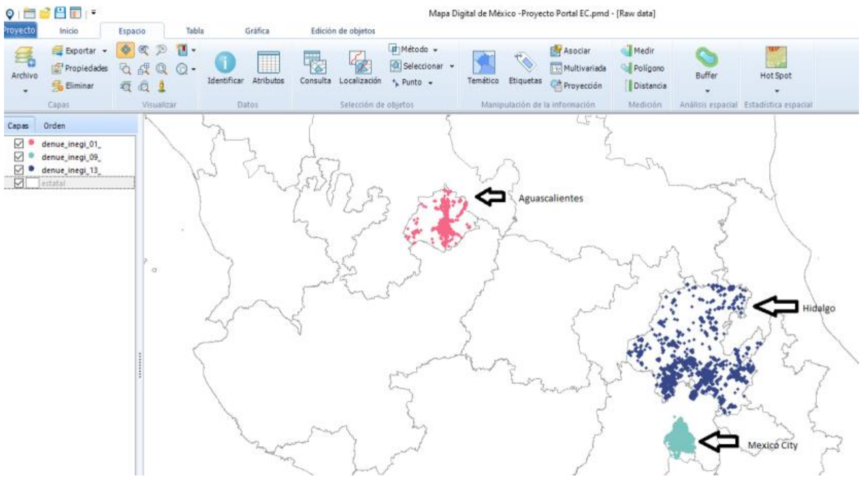

Fig. 5. Economic units in DENUE 2019 in Mexico City, Aguascalientes and Hidalgo.

14. Finally, the map created with the different layers was published as a Web Map so that anyone interested in finding a place or business that promotes any activity of the circular economy R-ladder can find its location.

4 Results

The original DENUE database containing approximately 675,261 economic units within the states of Mexico City, Hidalgo and Aguascalientes, was reduced (through the R-ladder filter classification) to less than 45,000 units involved in a circular economy activity. Table 3 shows the original number of economic units in each state (Fig. 5) and the numbers after the reclassification using the R-ladder.

There were some instances of ambiguity in classifying certain businesses. Some NAICS codes were difficult to reclassify because they comprised many kinds of companies and not necessarily all of them were related to circular economy activities. In the future, we want to automate this reclassification process using machine learning techniques, but at this point, all was done by human analysis.

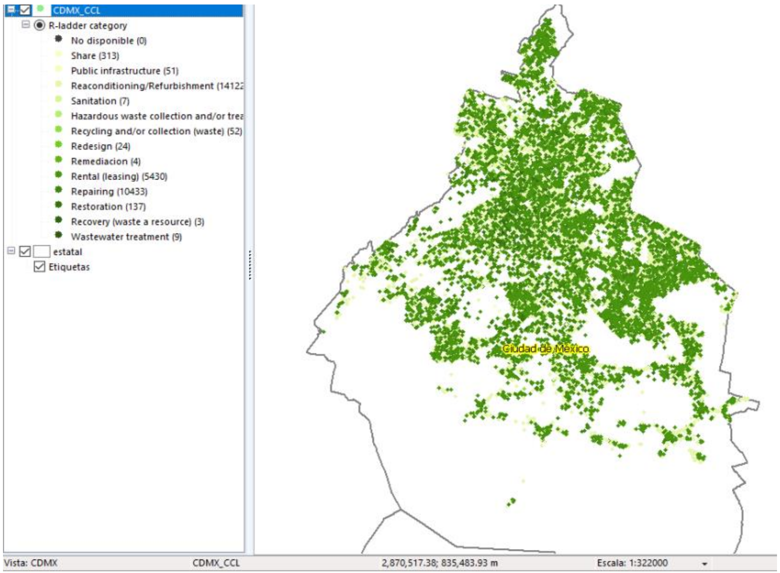

Fig. 6. Economic units with circular economy related activities in Mexico City.

We decided to keep some economic units entirely in a new category. For example, bioremediation service companies. Bioremediation would be a form of "recovery", but we decided to keep it in its own category and include it in the platform because of the relevance of these companies to preserving natural capital.

4.1 Reclassification of Economic Units (Micro, Small and Medium Businesses)

In Mexico City, according to the reclassification results, we found 30607 economic units that potentially do circular economy activities. They got subclassified like this (Fig. 6):

- 14122 are within the Reconditioning/Refurbishment category • 10433 are in the Repairing category,
- 10433 are in the Repairing category,
- 5430 Rental (loan),
- 313 Sharing,
- 137 Restoration,
- 52 Recycling and/or waste collection (micro and small places that receive materials such as plastics, paper, metal, etc., to sell them to bigger companies),
- 51 Public infrastructure,
- 24 Redesign,
- 22 Collection and treatment of hazardous waste,
- 9 Wastewater treatment companies,

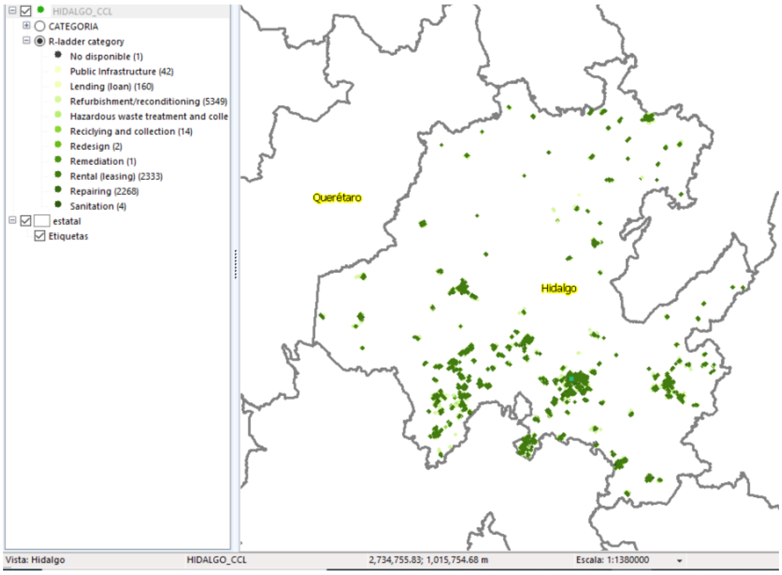

Fig. 7. Economic units with circular economy related activities in Hidalgo.

- 7 Sanitation companies,
- 4 are Remediation (wastewater or waste) companies and
- 3 are Recovery (waste as resource) consultancies.

As for the state of Hidalgo, the number of economic units within an R-ladder activity is 10175. At the same time, they are subclassified as follows (see Fig. 7).

- 5349 belong to the Refurbishment and/or Reconditioning category,
- 2268 belong to the Repairing category,
- 160 economic units are in the lending (loan) category (for example, book loans in public libraries). This category is not in the R-ladder, but it is considered a circular business model since it changes the focus from selling to lending something, thus contributing to the circular economy's second principle,
- 42 economic units are public infrastructure (such as water treatment plants, vehicle emissions check centers),
- 14 are Recycling/waste collection centers,
- 4 are Sanitation businesses,
- 2 are design companies (Redesign),
- 1 is a Remediation company and
- 1 is a hazardous waste treatment company.

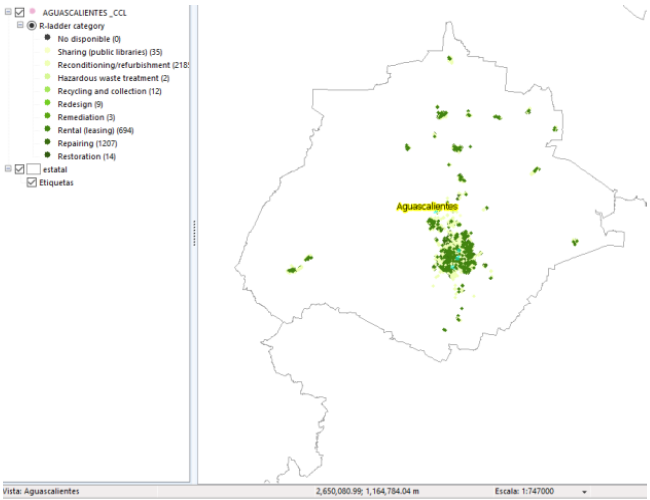

Fig. 8. Economic units with circular economy related activities in Aguascalientes.

Finally, for Aguascalientes we found 4161 economic units divided into the following categories (see Fig. 8).

– 2185 belong the reconditioning/refurbishment category reconditioning/refurbishment category,
– 1207 are repairing businesses,
– 695 are rental (leasing) companies,
– 35 are sharing entities (such as public libraries),
– 14 are within the restoration category,
– 12 are recycling and collection centers,
– 9 are redesign companies,
– 3 are remediation companies,
– 2 are hazardous waste treatment companie.

4.2 Circular Economy Web-Platform Prototype

The web map showing SMMEs with R-ladder activities is shown in Fig. 9.

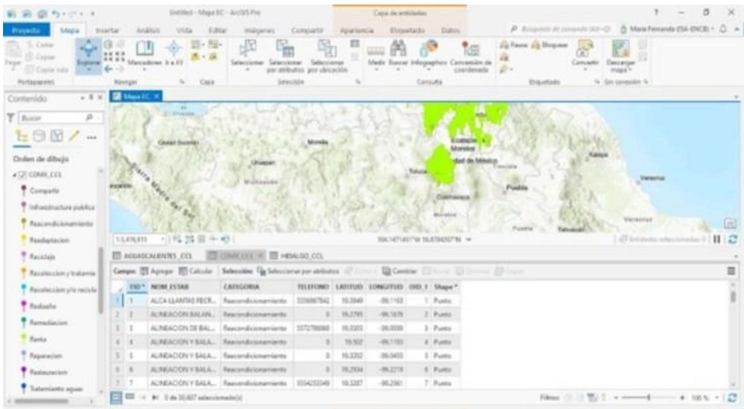

Fig. 9. Web-GIS layout using ArcGIS Pro.

Fig. 10. Points representing the economic units with R-ladder activities in Mexico City.

ArcGIS Pro and ArcGIS Online are partner platforms that enable users to create, explore and visualize georeferenced information. They also allow data from multiple sources to be integrated, edited and analyzed to perform spatial analyses among many other things. In addition, the ArcGIS suite enables the creation of geographic data platforms and websites without the hassle of computer programming. This is the reason this software was chosen for the creation of the circular economy platform. Arc GIS Online was used to create the web platform prototype.

Each state has a different number of economic units. However, since all the information provided by INEGI is already georeferenced and standardized to the Mexican norms to publish geographic information it was no problem at all to upload the data for the EC platform in the same standard.

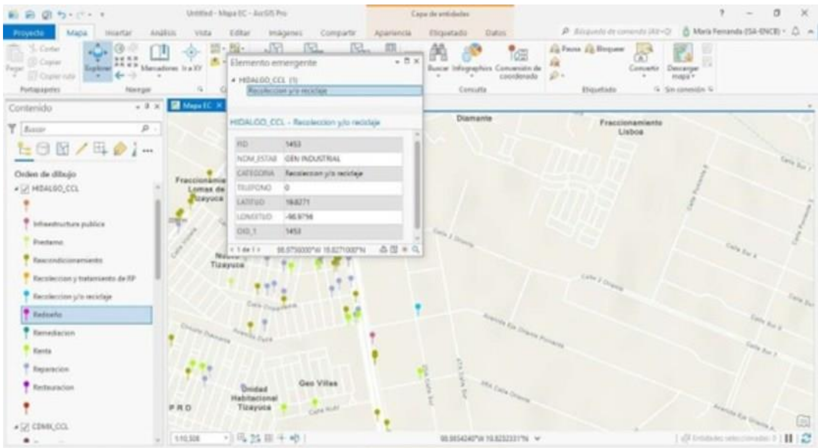

Fig. 11. Web-GIS platform prototype showing details of an economic unit in Mexico City.

A unique color for each R-ladder category was set for the three layers (CDMX, Hidalgo and Aguascalientes) to easily identify each R (see Fig. 10).

We are still working on improving the visualization feature to make it as user-friendly as possible. In another step of this project, we also aim to use the platform as an educational tool to explain what circular economy it is about and how it relates to the activities that many MSMEs do.

Before publishing the Web-map in ArcGIS Online, some points were randomly checked so that the descriptive table corresponds to the R-ladder category assigned (see Fig. 11)

Beyond visualization, this platform aims to provide quantitative and qualitative in-formation of companies engaged in circular economy activities and be a data analysis source of MSMEs involved in CE. According to Nery et al, digital technologies can be used to foster the circular transition by MSMEs [12] and MSMEs can accelerate the global green transition due to their proximity to the local environment and work force. [13]. However, while large enterprises and policymakers have made steps in adopting CE practices, small and medium-sized enterprises (SMEs) face unique challenges due to limited resources and expertise [14].

The web-GIS platform can be accessed in the following link: https://arcg.is/8THKP0.

5 Conclusions

According to the results, a total of 44,943 micro, small, and medium-sized companies across three Mexican states are involved in Circular Economy practices, although many remain unaware of their participation in this model. The platform aligns the eco-nomic units and their activities with the R-Ladder which at the same time is a hierarchical way to present the three principles of circular economy.

One of the main challenges of the reclassification step was to select DENUE codes (in NAICS standard) that relate to activities on the circular economy ladder and reclassify them into some R (repair, remanufacturing, recycling, etc.). Although there is room for further improvement the categories shown are true to their activity.

Efforts are underway to process and incorporate data from an additional 29 states into the Circular Economy Platform. Once the NAICS codes related to circular economy activities are identified, it is easier to filter them out in any other state database. Still, we are coming up with an algorithm to download and process the information automatically (from the DENUE website) so that it can be updated at least every year.

Completing and updating the GIS platform is crucial to release the platform to the public. The crowdsourcing tool is not ready yet, but its intention is to provoke public involvement in support of circular economy activities and, at the same time, be a way to keep the platform updated. We will make sure to implement a similar procedure as INEGI does. Since DENUE is under control of the Mexican Institute for Geography and Statistics and all the information is censed and updated using international standards, our data (DENUE) will also be reliable for next years.

References

1. Ellen MaCarthur Foundation: Circular Economy Introduction. EllenMaCartur Homepage. https://www.ellenmacarthurfoundation.org/topics/circular-economy-introduction/overview.
2. Díaz-Bone, H., von Zahn, F., Schreiner, A.: Circular Economy as a Cornerstone for Meeting the Goals of the Paris Agreement: A Roadmap Towards CE-Smart NDCs. GIZ (2021)
3. WDCD Homepage: How to Enter–WDCD. No Waste Challenge (2021) https://nowaste.whatdesigncando.com/about/how-to-enter/.
4. Araujo-Morera, J., Verdejo, R., López-Manchado, M.A.: Sustainable Mobility: The Route of Tires Through the Circular Economy Model. Waste Management, 126, pp. 309–322 (2021) doi: 10.1016/j.wasman.2021.03.025.
5. INEGI: Directorio Estadístico Nacional de Unidades Económicas 2020, Información para la actualización e incorporación de unidades económicas al DENUE; datos a noviembre de 2020 (2020) https://www.inegi.org.mx/rnm/index.php/catalog/587.
6. INEGI: Directorio Nacional de Unidades Económicas DENUE. Censos Económicos (2014) https://www.inegi.org.mx/app/mapa/denue/default.aspx.
7. U.S Government: United Stated Census Bureau (2023) https://www.census.gov/naics/.
8. INEGI: Síntesis metodológica del sistema de clasificación industrial de América del Norte: SCIAN 2013 (2013)
9. INEGI: Comunicado de prensa, 285(20), pp. 1–2 (2020)
10. SDGS UN: The 17 Goals Sustainable Development (2021) https://sdgs.un.org/goals.
11. Ma, H.W., Shih, H.C., Liao, M.I.: Circular Economy and New Research Directions in Sustainability. In: Chen, C., Chen, Y., Jayaraman, V.: Pursuing Sustainability. International Series in Operations Research & Management Science, 301 (2021) doi: 10.1007/978-3-030-58023-0_6.

12. Neri, A., Negri, M., Cagno, E.: The Role of Digital Technologies in Supporting the Implementation of Circular Economy Practices by Industrial Small and Medium Enterprises. Business Strategy and the Environment, 32(7), pp. 4693–4718 (2023) doi: 10.10 02/bse.3388.
13. Gennari, F.: The Transition Towards a Circular Economy. A framewok for SMEs. Journal of Management and Governance, 27(4), pp. 1423–1457 (2023) doi: 10.1007/s10997-022-09653-6.
14. Ahmadov, T., Durst, S.: SMEs on the Way to a Circular Economy: Insights from a Multiperspective Review. Management Review Quarterly (2023) doi: 10.1007/s11301-023-00380-2.

ISSN 1870-4069

Hazard Identification Procedure, Risk Analysis, and Determination of Controls for the Failure of a Line Valve in a Gas Pipeline

Vladimir Avalos-Bravo[1], Adriana Guzmán-Ruvalcaba[2], Patricia Alejandra Vallejo-Landa[3]

[1] Instituto Politécnico Nacional, CIITEC, SEPI-ESIQIE, UPIEM, SARACS, Mexico

[2] Instituto Politécnico Nacional, SEPI-ESIQIE, Mexico

[3] Instituto Politécnico Nacional, ESIQIE, Mexico

ravalos@ipn.mx, adrianagr.iqp@gmail.com, alejandra.landa@hotmail.com

Abstract. Petroleos Mexicanos (PEMEX) is a Mexican company that produces, transports, and refines petroleum products that cover the three segments of the oil value chain: downstream, midstream, and upstream. This company had around 8,400 kilometers of pipelines for the transportation of petroleum products and approximately 5,200 kilometers of oil pipelines for the transportation of crude oil. The operation of these transportation systems has an implicit risk that can affect the population and the environment when unwanted events occur such as leaks due to failures, mechanical integrity, or, where appropriate, leaks caused mostly by the illicit hydrocarbon market. In natural gas, on some occasions derived from a risk analysis performed during the design stage, a safe stop condition due to failure applied in the automated system of the gas ducts is necessary. For the example analyzed, humidity was found inside a safe box, it damaged the Limit Switch and caused a double-position failure, and it was not displayed in the SCADA system; some conclusions are that the failure report system activities are necessary, or anything related to the automated system. For all the parties involved, it is necessary to have a common agreement and criteria to take preventive and/or corrective actions in the gas duct transport management. Carrying out inspections regularly is needed because, for the analyzed case, the failure was a protection of the control system that sends a safe stop when detecting a disagreement in the feedback of the limit switch damaged by the presence of water.

Keywords: Risk analysis, pipelines, PLC, SCADA, control determination.

1 Introduction

One of the main activities in the entire value chain of the oil industry is the transportation of oil and its derivative products, which requires an infrastructure that can cover the entire national system and the demand for oil products. Gradually, the actors involved in fuel transportation and distribution logistics improve their options to cover increasingly larger market areas.

To bring products to consumer markets, the main means of transportation are pipelines, tankers, tank cars (railway), and tank cars (pipes). Strategic considerations for investments in fuel logistics infrastructure are based on a thorough analysis of the operating conditions of the means of transportation and distribution, and the growth of local markets. For local fuel distribution or short distances, self-tankers are used; these carry the fuel from the Storage and Distribution Terminals (SDT) to the service stations. Despite the different standards and inspections that are carried out continuously throughout the value chain, it is important to mention that since the oil industry is one of the riskiest due to the nature of hydrocarbons, the analyses and evaluations Risk assessments are carried out to find the causes and consequences of incidents and accidents.

1.1 PEMEX Logistica

Logistics involves managing order processing, inventory, transportation, and combining warehousing, material handling, and packaging, all of this integrated through the business network. The goal of logistics is to support the operational requirements of customer procurement, manufacturing, and supply. The challenge within a company is to coordinate functional capacity into an integrated operation that focuses on serving customers [1].

To supply the population in the national territory, the oil products produced or imported are transported from their point of origin, which may be a refinery, a maritime terminal (MT), or an entry point, to the storage terminals and distribution (SDT) supplied by service stations. This involves the passage between stations or plants for com-pression, pumping, processing, storage, and finally distribution. Pemex Logistics has six systems distributed in the eight statistical regions of the country [2].

The main pipeline builder throughout the country has been PEMEX. It has pipelines for the collection of crude oil and gas extraction from extraction wells, transportation to refineries, petrochemicals, and gas processing complexes, and the distribution of final products to SDT and final consumers (service stations) [3]. The oil distribution infrastructure can be analyzed from two areas: storage and transportation. Storage can be defined as the activity that includes the reception of petroleum products, property of third parties, at the reception points of their installation or system, keeping them in deposit, safeguarding them, and returning them to the depositor or whoever he or she designates. As well as, transportation is the activity of receiving, delivering, and, where appropriate, driving oil products from one place to another through pipelines

Table 1. PEMEX Logistica infrastructure [4].

Pipelines products	Kms	Facilities	Number	Transport Units	Number
Polyduct	8,390	Storage and dispatch terminal	74	Tanker	16
Oil duct	6,291				
LPG duct	1,395				
Gas duct	1,246	Liquefied gas distribution terminal	10	Pipe Tank	1,485
Chemical duct	828				
Gasoline	381				
Fuel oil	145	Navy terminal	6	Tank car	520
Basic petrochemicals	1,294				
Turbosin duct	81				

or other means [2]. Within the oil value chain, transportation is one of the most important activities.

The infrastructure is composed of pipelines, facilities, and transportation equipment as shown in Table 1 which are distributed in the various regions of the country.

1.2 Pipelines

These facilities, unlike others, are not located within an industrial complex with security features; on the contrary, they are located throughout lands owned by third parties, between cities and roads, or on agricultural lands, rivers, and natural landscapes. Which indicates a high probability of leaks or spills, contamination, and explosions [5]. We can classify the pipelines by oil pipelines, which are those that transport crude oil from the extraction areas to the refineries, petrochemical, and gas processing complexes, and the polyducts that transport all types of already processed fuels, mainly gasoline, and diesel [5].

The South-Gulf-Central-West Zone system is the largest, with a length of 4,962 kilometers, which allows the flow of oil products from the Gulf of Mexico to the center of the country and the Bajio, as well as to the Pacific through the Isthmus of Tehuantepec. The second is the Zona Norte system, with 3,152 kilometers, which has three cross border pipelines with the state of Texas in the United States, for the import of fuels. The remaining systems total 770 kilometers, between Guaymas, Rosarito, Topolo-bampo and Progreso.

The main hydrocarbons transported in Mexico are gasoline, diesel, jet fuel and fuel oil with 400.3, 273.8, 56.3 and 307.5 thousand barrels per day generated respectively, according to the Pemex Statistical Yearbook. In addition, other petroleum derivatives are generated in smaller quantities (68.4 thousand barrels per day in total) [6].

Gas pipelines are used to transport fuel gases on a large scale. Its function in current economic activity is very important. It consists of a conduction of steel pipes, through which the gas circulates at high pressure, from the place of origin. They are built buried in trenches at a usual depth of 1 meter. Exceptionally, they are built on the surface [7].

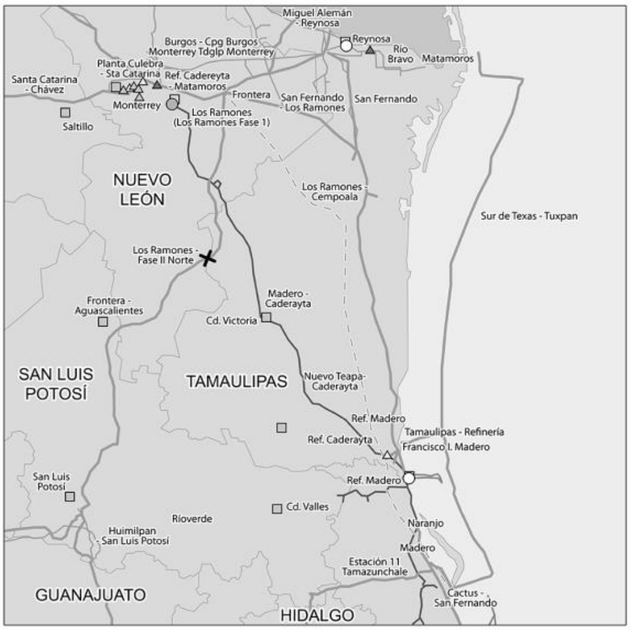

Fig. 1. Gas pipeline as a case study [5].

For the purposes established in the Hydrocarbon Sector Law, the following types of gas pipelines are considered:

1. High-pressure primary natural gas transportation pipelines: Those whose maximum design pressure is equal to or greater than 60 bars.

2. Secondary transport gas pipelines: Those whose maximum design pressure is between 60 and 16 bars.

3. Distribution gas pipelines: Those whose maximum design pressure is equal to or less than 16 bars and those others that, regardless of their maximum design pressure, are intended to convey gas to a single consumer starting from a gas pipeline of the Basic or secondary transport network.

For this study, a section of gas pipeline located on the border of Nuevo León and Tamaulipas was taken into consideration.

The Los Ramones project, which is the one that distributes gas from the coasts of the Gulf of Mexico and goes up to Hermosillo, Sonora, is one of the most important. from the northern zone and that are managed by foreign companies.

The gas pipeline runs from Camargo (Tamaulipas state) to Apaseo el Alto (Guanajuato state). Its route passes through key crossroads in Los Ramones (Nuevo León state), Villagran (Tamaulipas state), and Villa Hidalgo (San Luis Potosí state) before continuing south through San Luis de la Paz (San Luis state). Luis Potosí) and the Querétaro Industrial Park (Querétaro state) to Apaseo el Alto.

1.3 Chemical Accidents and Risk Analysis

Throughout history and the technological development that humanity has experienced, chemical substances have been present, whether in their liquid, solid, or gaseous state; At first, all these substances were used for a series of fortuitous discoveries and later to generate new technological developments that improved the quality of life of society. But the manipulation of all these chemical substances was accompanied by the first fires and explosions of which there is a record, situations that were repeated again and again, until the different physical and chemical properties of the substances used in the processes, the criteria that have been taken as a basis to make a classification of the most critical technological accidents worldwide are based on the number of people killed or affected by this type of events, the ecosystems that are damaged or the adverse impact that they cause. cause subsoil pollution in rivers, lakes, or seas and an increase in pollutants in the environment [8].

Due to the large number of accidents that have occurred in the world in the hydrocarbon sector, it is necessary to carry out a risk analysis. This can be done with different degrees of detail, depending on the risk, the purpose of the analysis and information, as well as the data and resources available. The analysis can be qualitative, semi-quantitative quantitative, or a combination of the three cases, depending on the circumstances. Consequences and their likelihood can be determined by modeling the outcomes of an event or set of events, or by extrapolating from experimental studies or available data. Consequences can be expressed in terms of tangible or intangible impacts, in some cases, more than one numerical value or descriptor is required to specify the consequences and their possibility for different times, places, groups, or situations [9]. Decisions should consider the broader context of the risk and include consideration of the tolerance of the risk by other parts of the organization that benefit from the risk. Decisions should be made in accordance with legal, regulatory, and other requirements. In some circumstances, risk assessment may lead to a decision to conduct further analysis. The risk assessment may also lead to a decision not to address the risk in any other way than by maintaining existing controls.

This decision will be influenced by the organization's attitude toward risk and the risk criteria that have been established [10].

The problem that will be analyzed in this paper is the failure in a section of a gas pipeline called Los Ramones. It is reported by the personnel that manages it that in the SCADA system, a partial closure signal is emitted in a mainline valve simultaneously. 95%, for this reason, and given that the closure cannot be carried out remotely, people from the different profiles assigned to the administration of the gas pipeline are summoned to carry out a risk analysis on the reported situation and avoid a major disaster.

2 Methodology Application to the Study Case

To carry out the proposed methodology, it is essential to collect all data regarding the reported failure. The Activity Manager must consult any matrix that has been

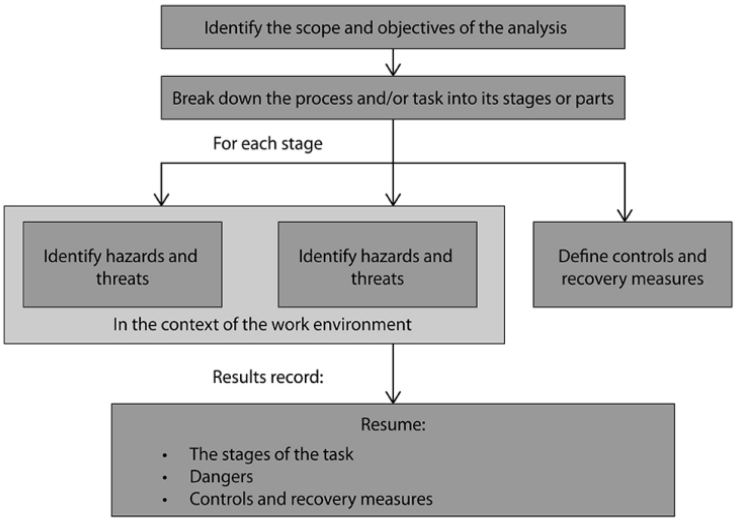

Fig. 2. Proposed Risk Analysis Flowchart.

prepared to show the library and/or inventory of the main safety and environmental risks and controls identified in the global process and must subsequently carry out the specific risk analysis (proposed format) for the task, whether routine or nonroutine, to detail the controls that will be implemented in the specific activities to be carried out.

The global risk analyses of the process (such as HAZOP studies, or any other risk matrix) and the task-specific risk analysis (prepared format) must be prepared based on the following criteria:

The analysis of all main stages of the process (gas transportation, operation, maintenance, dismantling, abandonment, etc.), and each step of the specific task, should be considered.

Activities that have the potential to cause significant incidents. Activities that have a history of incidents and accidents.

New activities or tasks that introduce new work methods.

The design of work areas, processes, use of machinery/equipment, type of operations, personnel requirements, requirements, and controls for contractors, subcontractors, service providers, and suppliers must be considered, and adaptation measures for personnel with special needs must be included. when it's requested.

Situations that are not controlled by the organization and that occur outside the workplace, but that may cause injury or damage to the health of people in the workplace and damage to the environment, for example: adverse weather, sabotage, crime, terrorism, outside personnel (visitors, contractors, suppliers, etc., who may cause damage), external factors such as damaged pipelines, nearby fires, road blockades, etc.

Physical capabilities and other human factors, for example: workloads, uncomfortable postures, fatigue, stress, anxiety, repetitiveness, skills, ergonomics, breaks, adaptation for personnel with special needs, etc.

	Severity	X (per)		Probability
1	No injury or First Aid, lost production time and negligible property damage, surface contamination.		1	Not likely to occur when safety or prevention measures are effectively implemented
2	Recordable Accidents with lost days, considerable property damage, plant stoppage of less than a week, moderate contamination.		2	Possible, it could happen at some point.
3	Fatality or Permanent Partial Disability, plant stoppage with multiple days, high property damage, contamination that exceeds regulatory limits.		3	Probable, probably has occurred several times or has already occurred on site.

Fig. 3. Approximated Risk Value.

Activities carried out under unusual circumstances.

One-time tasks.

As verification of the hazard management elements of an established procedure or work instruction.

The diagram shows the steps to follow for the procedure described above. To begin the risk analysis, it is necessary to determine the process and/or activity to be analyzed and establish the scope of the work and the objectives of the analysis. This scope must include the stage of the process, activity, or task that will be analyzed, and we must consider the possible effects of the environment in which the activity will be carried out, considering infrastructure, equipment, materials, substances, physical conditions of the place, etc.

Later we will separate the process into basic stages or steps describing the activities to be carried out and their sequence. We must detail each of these and define each change that will be made to mitigate the effects of the risks they may generate.

The description of each stage must provide a declaration of the activities to be carried out, begin each action in the present, and end with the subject that is being analyzed, be it a pump, a tower, valve, and omit references to dangers.

To identify hazards, it is necessary to examine each stage and consider the spatial context that could increase the hazard and any incidents that could result from the location in which the activities are carried out. It is very important to consider any incident that has happened previously and is related to similar activities to avoid its recurrence.

For the risk evaluation part, it is important to mention that this is an approximate value that is obtained by multiplying the severity X the probability according to the data in the format shown below.

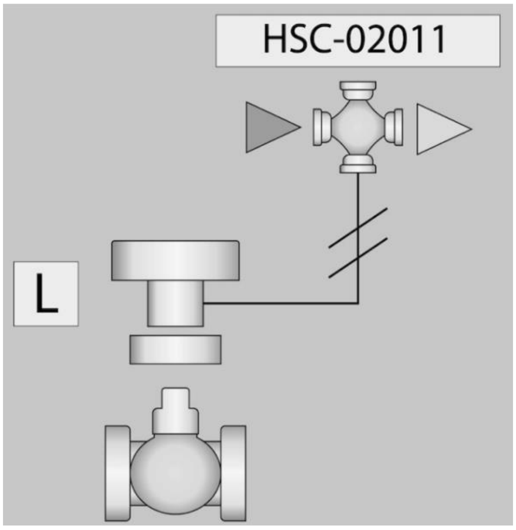

Fig. 4. MLV visualization in SCADA.

3 Discussion and Results

For 8 months there were no activities related to the Line Valve (MLV) itself, however, for this analysis, it is considered that the valve failed, and the SCADA technician visualized the valve in closed mode, and he informed the pipeline leader. about the status of the valve. The pipeline leader activates his on-call technical personnel to review the status conditions of the notified valve, so the SCADA leader receives a report that the valve is in closed status and informs the Operations Manager and the Pipeline Leader. The pipeline technician goes to the site to review the conditions of the Valve and performs valve opening maneuvers manually, leaving the 14" bypass open. Subsequently, they verify the current conditions of the valve with SCADA and report a failed status.

The SCADA Technician verifies the valve position via TVCC and notifies the pipeline technician of a second 50% closure on the valve.

The pipeline technician verifies and confirms to proceed to open the valve manually. He remains on site to monitor current conditions and after approximately 10 minutes he closes at 50% (for the third time). The pipeline technician opens the valve manually and tests the system by requesting the SCADA technician to close the valve and the closing command does not reach the valve, so the selector is positioned remotely to mode local.

The conditions that the open position remains are monitored again and in SCADA the visual mode sends the FAULT signal. The technician receives authorization to proceed to block the gas supply to the valve actuator from the pipeline leader, SCADA leader, and Operations Manager.

Fig. 5. Valve opening and visualization.

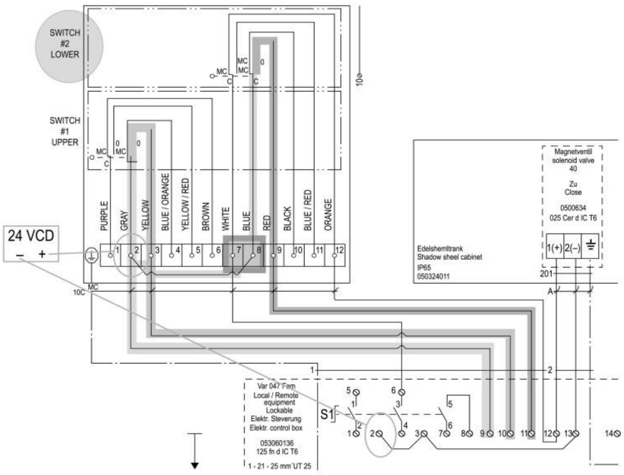

Fig. 6. Valve status before and after the error reported by the SCADA system.

3.1 Review and Inspection

Specialized personnel came to the site and verified that the tubing was not saturated/clogged by dust and that all the conduits were clean and free of blockages.

The Limit Switch Box was checked, and it was found that the closing screws on the top cover were loose, which were removed along with the cover, and it was observed that there was water inside the box.

The inside of the box was drained and cleaned, then the selector was changed from local to remote to verify any changes with the SCADA technician and the FAILURE mode continued.

Table 2. Research results and analysis. Part 1.

Cause	Recommendations	Reason	Actions to take
Due to humidity, it generates a bridge between lines 7 and 8 that sends a closing command.	Change the box of the Limit Switch Box.	By recommendation of the manufacturer and to guarantee the useful lifetime	Complete box replacement Field review of all valve locations with technical personnel
	Check that the hardware on all other boxes is adjusted correctly.	To validate that everything is in correct closure.	
	Arrange a visit with the maintenance personnel to carry out an internal inspection of the Limit Switch Box of all the valves of the gas pipeline systems.	To validate through reliability the findings generated in the internal inspection.	
Due to the tests carried out and having to repeatedly remove the lid.	Previously analyze the activities to be carried out and the implicit risks.	To minimize human factors of repetitions.	Performing task security analysis
	Present and continuous supervision whenever there are repetitive activities.	To guarantee the supplier's and operations recommendations comply with the procedure Performing the security analysis of the task	Always request the documents associated with the activity to be carried out (operational procedures, technical sheets, AST, Standards, etc.)
	To minimize human factors of repetitions.		

A local test of closing and opening the valve is carried out (after authorization by telephone with the SCADA technician) and the status is corroborated with the SCADA technician and his confirmation was that the FAILURE is maintained.

The valve is left blocked, gas is supplied to the valve actuator and specialized personnel are present to check the conditions of the valve.

The conditions of the valve and the position indicator are visually checked, subsequently, the conditions of the connections within the position indicator are verified.

Continuity tests were carried out based on the electrical diagram, which was provided by the manufacturer, and through these tests a damaged switch was detected, voltage measurement was also carried out and it was found that a CLOSED-OPEN signal was sent at the same time. It should be noted that the logic of a PLC sends a failure notification when the valve is like this.

Table 2. Research results and analysis. Part 2.

Due to bad practice of the contractor and client	Generate inspection procedures and formats for the delivery and reception of equipment, machinery, and facilities, among others.	To have transparently what is declared against what is received	At worktables with administration, supplies, and operations staff.
	Apply the controls established for DELIVERY-RECEPTION To have transparently what is declared against what is received	To have all parties involved in a common agreement	At a meeting with the people responsible for the DELIVERY-RECEPTION process
There are no visual or specific inspection recommendations from the manufacturer or supplier	Request the supplier/manufacturer for the technical sheet, valve inspection forms, operation manual	To objectively have all the points or aspects that must be reviewed preventively	Through the document/format that the manufacturer provides
Analyze possible failures during the delivery period on the day of the event. Request SCADA to review its history if any failure occurs in this period.	Request SCADA to review its history if any failure occurs in this period	To evaluate the conditions or possible causes that caused it	In previous review with a worktable
Limit Switch Box Diagnosis	Ask the manufacturer to review and diagnose the box	To determine the causes that caused the damage to the box and its relation-ship with the valve closure	Disassembling, reviewing electronic diagram, photographic report, final failure report
Origin of 24 VDC	Perform inspection on OC and port card	Find the signal that could activate the Limit Switch Box solenoid.	Simulating the conditions (bypassing pins 7 and 8) of the Limit Switch Box diagram

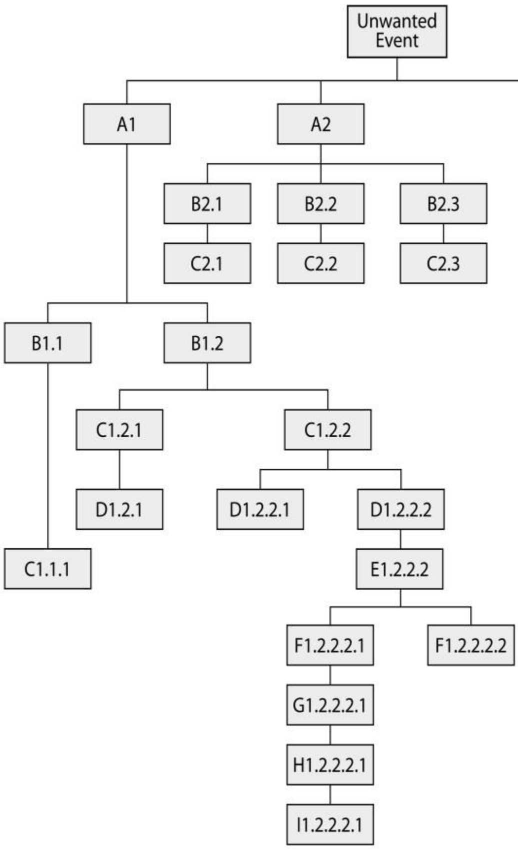

Fig. 7. Carried out analysis chart.

The SCADA technician is asked for the closing command to verify that the signal is arriving within the position indicator. This is confirmed on-site by the manufacturer's personnel.

For the coil to be energized, two conditions must be met:

– The remote local position switch must be set to remote.

– The position indicator should read OPEN.

Due to the data of the readings obtained in the SCADA system, it is possible to see that the valve is not closed due to the factors previously described, for this reason the manual opening was carried out to verify the way in which the SCADA system reflects it.

Due to the data of the readings obtained in the SCADA system, it is possible to see that the valve is not closed due to the factors previously described, for this reason, the manual opening was carried out to verify the way in which the SCADA system reflects it.

3.2 Unwanted Event Valve Closure at 95%

- A1 Receives closing command.
- A2 LINEBRAKE was activated. A3 By closing locally.
- A4 Closes manually
- B1.1 SCADA system sends the closing command. B1.2 By voltage source that activates.
- B2.1 Natural gas leak.
- B2.2 Due to obstruction in the supply tubing B2.3 By pressure difference.
- B3.1 Collaborator performs closing maneuver.
- B4.1 At the time of closure there was no presence of personnel inside the MLV. C1.1 Not powered.
- C1.2.1 The fuse does not work.
- C1.2.2 Within the limit switch box there are 2 points with voltage present (1, and 8) C2.1 Condition not met.
- C2.2 Condition not met.
- C2.3 The pressure differential condition was not met to generate the activation of the Line Brake.
- C3.1 Personnel does not carry out closing operations, the conditions are not met. D1.2.1 Because it is an independent system of these electrical terminals of the Limit Switch.
- D1.2.2.1 By manual manipulation and creating a bridge between 7 and 8.
- D1.2.2.2 Due to humidity it generates a bridge between 7 and 8 that sends a closing command.
- E1.2.2.2 Poor closing of the limit switch cover (loose screws) F1.2.2.2.1 Failure to correctly close the LSB cover.
- F1.2.2.2.2 Due to the tests carried out and having to repeatedly remove the cover. G1.2.2.2.1 The MLV DELIVERY-RECEPTION procedure was not complied with. H1.2.2.2.1 There is no procedure.
- I1.2.2.2.1. Due to the bad practice of the contractor and client.

A fault tree analysis was applied to list the possible causes of the valve failure and determine those that occurred and led to the unwanted event reported by the SCADA system.

4 Conclusions

As a recommendation, you should continue with the reports and reports of failures in this type of pumping system in the midstream segment, as well as activities or things related to the gas pipeline and all the parties involved to have a common agreement and criteria to consider preventive and/or corrective actions.

The operations department of any pipeline monitoring area must carry out an inspection as a visual methodology for its field tours and continuous maintenance of the MLV's equipment.

The failure apparently is a protection of the control system that sends a safe stop when detecting a disagreement in the feedback of the limit switch damaged by the presence of water.

In the case of natural gas, on some occasions, the programmers, and operators, derived from the risk analysis carried out in this work and during the design, must have set a safe stop condition due to failure, therefore, when the operating personnel goes to open the valve, the PLC automatically orders it to close.

The solution to this type of problem is to have a maintenance mode in the control logic to allow people to operate in manual mode since the humidity found inside the box damaged the Limit Switch, which caused the double position failure. displayed in the SCADA system.

The supplier issued the recommendation that the starting system should have been completely changed in the box to put the Valve into operation and continue with the analysis of the origin of the failure.

The Supplier carried out the inspection of the box and all its internal components to determine if the failure caused by humidity was related to the closure of the valve, but the operational work group carried out some additional inspections as support to reinforce the diagnosis issued by the from the supplier and found no evidence of any other damage.

Acknowledgments. This project was funded under the following grants: SIP-IPN: No 20230046 with the support of CIITEC-IPN, and special thanks to Hernan Abraham Perez Flores for the figure's edition.

References

1. Bowersox, D., Closs, D., Cooper, B.: Administración y logística en la cadena de suministros. McGraw-Hill Interamericana (2007)
2. SENER.: Diagnóstico de la industria de petrolíferos en México (2018) https://www.gob.mx/sener/articulos/diagnostico-de-la-industria-de-petroliferos-en-mexico.
3. Novedades Tabasco: Humedales sobre "Bomba de Tiempo" (2018) https://novedadesdetabasco.com.mx/2018/08/29/humedales-sobre-bomba-de-tiempo/.
4. PEMEX.: Almacenamiento. Ciudad de México, México: Petróleos Mexicanos (2019) https://www.pemex.com/nuestro-negocio/logistica/almacenamiento/Paginas/default.aspx/.
5. Llano, M., Flores, C.: Ductos, ¿por dónde circulan lo hidrocarburos en México?. Heinrich Böll (2017) https://mx.boell.org/es/2017/07/26/ductos-por-donde-circulan-los-hidrocarburos-en-mexico.
6. Arvizu, L.: Propuesta para el manejo de materiales y residuos peligrosos dentro de una estación de lavado y control de remanentes de autotanques que transportan hidrocarburos. Tesis de maestría UNAM (2014)
7. Eisenberg Gas Service: ¿Que son los gasoductos? (2023) https://www.eisenberg.com.mx/infraestructura/gasoductos/.

8. Vilchis, L.: Análisis HazOp en un reactor intermitente como modelo preventivo de riesgo industrial (2020) https://tesis.ipn.mx/handle/123456789/28525/.
9. Barragán, B.: Análisis y evaluación de Tomas Clandestinas en México. Caso: Tlahuelilpan Hidalgo (2020)
10. Norma ISO-31000: Sistema de Gestión de riesgo y seguridad (2022) isotools.org/normas/riesgos-y-seguridad/iso-31000.

ISSN 1870-4069

Application of Machine Learning in the Development of Mobile Applications for Geographic Information Systems Using CreateML, CoreML and Mapkit for iOS

Eduardo Eloy Loza Pacheco, Mayra Lorena Díaz Sosa,
Christian Carlos Delgado Elizondo, Mayra Olguin Rosas

Universidad Nacional Autónoma de México,
Mexico

{eduardo.loza, mlds, ccdelgado}@acatlan.unam.mx

Abstract. Applications of Geographic Information Systems (GIS) are found in Urban Planning, meteorology, Biology, Data science, etc. GIS is implemented in several technologies such as desktop, mobile, etc. Mobile applications are now important because of the mobility, and connectivity they offer. More GIS applications are now migrating to a mobility strategy. Additionally, there is a wide set of technologies we can use in the development of mobile applications for example for Operative systems like Android OS and iOS. In the case of Apple, it has developed an interesting ecosystem of Technologies such as ARKit for Augmented Reality, Virtual Reality that has become more popular with the Introduction of Apple Vision and the visionOS, Technologies for Artificial Intelligence like Create ML and CoreML, Map Kit is designed for the use of maps and its manipulation. The technologies behind these developments are. Swift UI and UIKit. Finally, the objective of this work is to describe a Framework to use technologies to create smart GIS.

Keywords: Machine learning, artificial intelligence, core ML, create ML, iOS.

1 Introduction

The use of Geographic Information System is used in every field of knowledge. For example, in [1], they mention that GIS is no longer solely associated with geography. Today we can find applications to GIS in social studies, artificial intelligence for example in spatial and qualitative reasoning, urban studies, and education, among other applications. On the other hand, technologies are become more sophisticated and miniaturized. We can have all the characteristics of a computer machine in our hands. One example is smart phones. In addition, there are a wide variety of applications that use components of a GIS. For example, the application Waze or Google Maps can be used to describe a route. Uber application is used for transportation of people, food, or delivery.

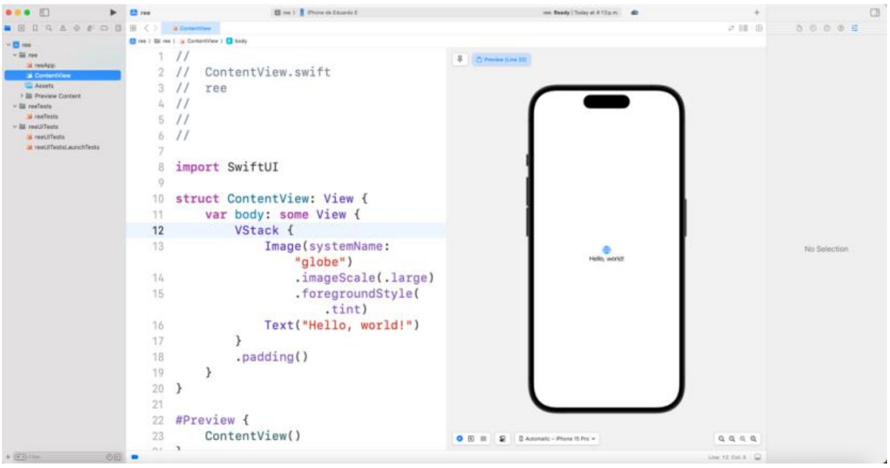

Fig. 1. XCode Technologies and Swift UI platform.

Mobile applications are now a growing environment where enterprises, organizations, and institutions can develop a solution or application to be more competitive. In the case of Apple, they develop a SDK (Software Development Kit). In the SDK developers can build useful, native applications. These libraries are Core ML and the set of applications that are Create ML and CoreML [3]. Core ML is a Machine Learning tool that allows training the machine to obtain a model that can be used with a specific purpose. Anode library is MapKit with a set of functionalities to describe maps and manipulate them [4]. XCode is a software where we can implement mobile applications using Swift Programming Language and Swift UI framework.

2 Related Work

Today the use of Geographic Information System (GIS) is beyond geography. The popularity of GIS is increasing thanks to its utility. Now we found GIS in social sciences, health, crime, climate change, remote sensing [5], etc. GIS is almost ubiquitous in every science field [1]. On the other hand, applications related to Artificial Intelligence are very common in literature, for example in [6], GIS and Machine learning are used to map minerals. The objective is to find potential targets where the mineral can be extracted (The article mentions some zones of China). In [7] artificial intelligence is used to simulate avalanches to prevent tragedies. Finally, [8] shows the importance of preventing vegetation vulnerability by monitoring and predicting.

On the other hand, the rise of mobile applications allows to implementation of a variety of applications in a smartphone. For example, in [9] mobile applications are useful in the event of disasters. In [10] the work shows the integration of GIS into mobile applications. Finally, as we see, there is an evolution of GIS applications to mobile applications, sometime because their portability and computer power.

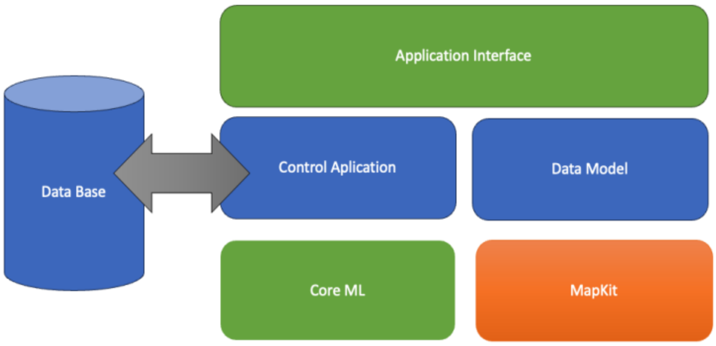

Fig. 2. Proposed architecture.

3 Proposed Architecture

The proposed Architecture is showed in fig 2. The architecture is based in Vista Controller Model where elements are divided in areas of functionality. The first layer is recommended to design the application interface, and the functionality of the buttons. Then we define the application control where we model the communication between the Database, and the components necessary of the GIS Application, for example if the applications need an implementation of Map GIS to show data on a machine learning model whether an artificial intelligence is needed. Finally, we require a block where we can model the data of the application.

3.1. Application Interface

The application interface is defined according to the design proposed in [11]. Where the process of designing and panning is divided in four stages: Brainstorm, Plan, Prototype and Evaluate. Fig 3 shows the cycle of development of a mobile Application.

– Brainstorming. Allows us to identify different problems and propose a solution. This stage prepares us to configure the elements that an iOS application needs to address.
– Planning. Is the stage where we can describe step to step the details of the solution and the means to achieve a goal.
– Prototyping. In this stage, we create a visual example of the application. Usually to our users. The idea is to receive feedback from our users and all the stakeholders in the project.
– Evaluating. Allow us to present the applications to several users in order to find possible errors.

Fig. 3. The design Cycle taken from [11].

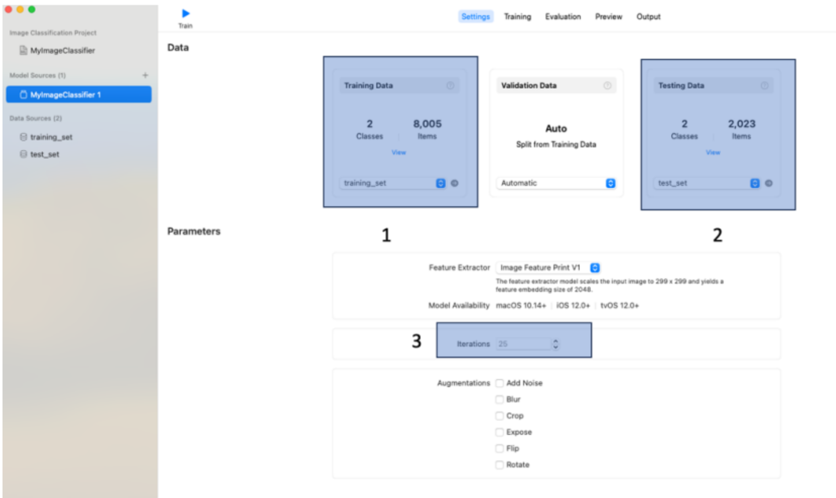

Fig. 4. Create Model window, section of training.

3.2. Control Application

Core ML is a technology that allows to produce machine learning models to process images, and video, capture motion, identify sounds, text recognition. To create a model it is necessary to use Create ML. Create ML is a tool that permits to generate models for hand pose classification, object detection, image classification, text classification, word lables among other applications.

Later Create ML generate a model (with ml file extension) that can be added to the main application in XCode. Suppose for example we would like to identify cats from dogs, in figure 4 we can see the main window where in section 1 is added all the images necessary to train the model. In section 2 it is necessary to add the image required for the training stage (Note the validation data is set in auto). Finally, in section number three we specify the number of epochs required to train the model and minimize the error.

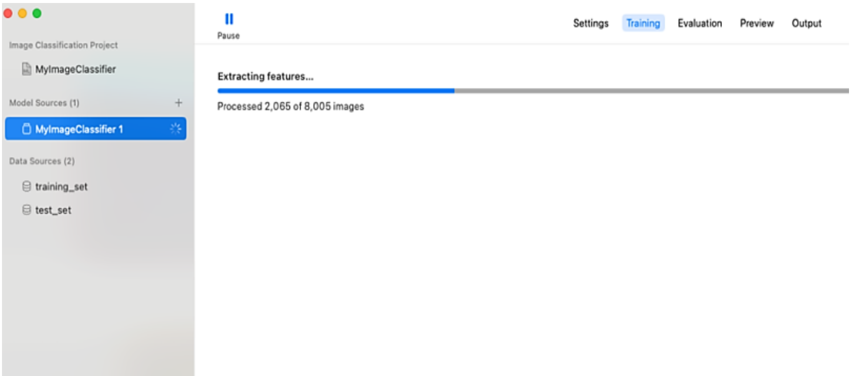

Fig. 5. Training process of the model.

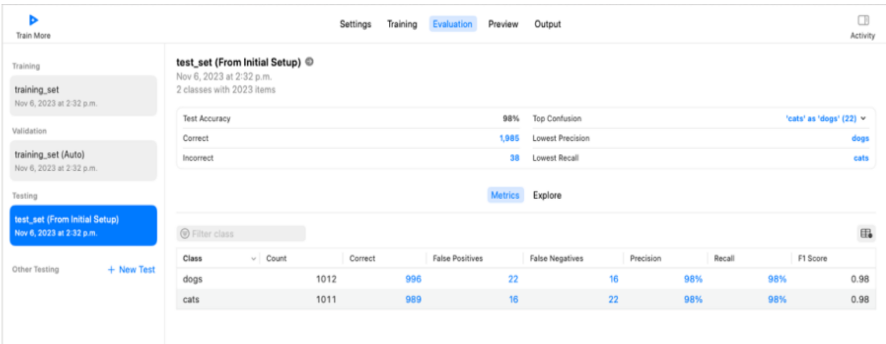

Fig. 6. Results of the training process.

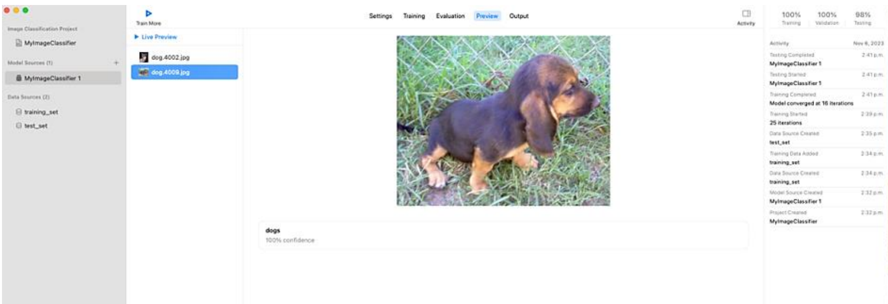

Fig. 7. Preview mode.

Later in Fig 5, We can observe the training process. It is worth mentioning that thanks to Apple technology and the set of processors M2 [12], it is possible to train a model using a very reduced amount of time. Then in Figure 6 we observe the results of the training model. We can see important information such as the accuracy of the model and other important metrics such as recall and F1 score. Finally, in figure 7 we can add any image and test the model. Finally, we can export the model to add it to our Project folder in XCode.

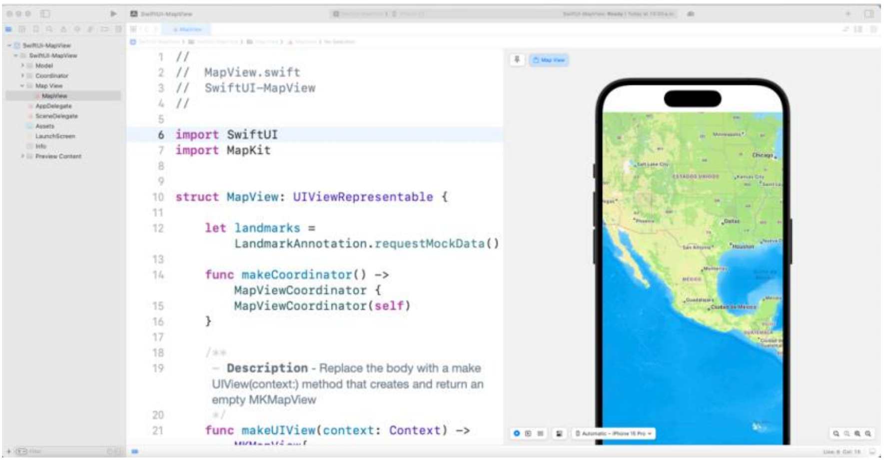

Fig. 8. Mapkit Framework.

3.3. MapKit

Finally, Mapkit is a technology that can accept structure types like Jason, GEO Jason in the case of spatial. Additionally, can detect the geoposition. Mapkit framework also permits adding maps, annotation, responding to user interactions, providing text completion [4], and now is integrating with ESRI [13].

3.4. Integration with GIS

GIS is important in a variety of science fields, for example Economy, Meteorology, Civil Engineering, Architecture, Computer Science, Social Studies, etc. According to [Introduction to GIS], the key components of GIS are:

Software: Necessary to implement algorithms, tools, methods, etc. The development of this branch are need and more types of software from standalone applications, web applications, and now mobile applications. Additionally, the present requirements need to add elements of artificial intelligence, computer graphics, etc.

- Data: Data is the most important feature of this century. Now data drives the operation [14]. Data needs to have a format and require to be complete, correct, and precise. Sometimes need to be preprocessed to be scaled or transformed.
- Protocols: Protocols are necessary in GIS to describe the data entry and updating.
- Hardware: It is necessary to consider the magnitude of the applications, according to it the use of memory, and interfaces.
- People: GIS identified different types of users, and experts, with no knowledge.

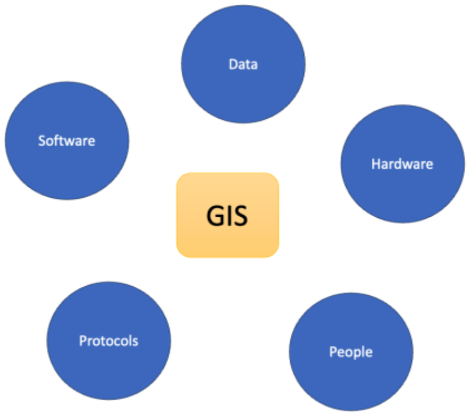

Fig. 9. Key Components of GIS.

Finally, to resume the component software is constantly updating. So it is necessary to implement ad hoc technologies and methodologies. In this case, ways to implement GIS in mobile applications using iOS technologies. In Addition to that Data needs to have a data model ad hoc to the iOS applications format. Protocols need spatial implementation according to the design of the application. The hardware needs to be implemented in iOS using native programming because of the particularities of the Apple Hardware. Finally, is necessary to have developers specialize in Swift, Objective-C and XCode.

4 Conclusion

The development of a mobile application needs to be measured in weeks, or months. The development time is getting shorter every day. Technologies needed today to build an application are becoming more sophisticated. It is important to have prebuilt frameworks, and libraries to develop in time.

Additionally, it is important to have an architecture to integrate all different technologies to identify different functionalities an application needs to implement. For example, a stage where the design of the application needs to be separated from the technologies that can be implemented for example Machine learning (MapKit) and display maps and different elements. MapKit is a robust technology developed by Apple that allows the addition of different functionalities a GIS has. On the other hand, machine learning is a recurrent topic that is often implemented in all kinds of applications. Create ML and CoreML help us to add interesting applications the application needs to implement.

Acknoweldgments. The authors would like to thank to Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Acatlán and DGAPA-UNAM (PAPIME PE305523 and PAPIME PE 301521) for the funds to support this work.

References

1. Ballas, D., Clarke, G., Franklin, R.S.: An Introduction to GIS (2017) doi: 10.4324/9781315759326-1.
2. Goodchild, M.F.: Giscience, Geography, Form, and Process. Annals of the Association of American Geographers, 94(4), pp. 709–714 (2004) doi: 10.1111/j.1467-8306.2004.00424.x.
3. Apple: Core ML. Apple Developer Documentation (2022) https://developer.apple.com/documentation/coreml.
4. Apple: MapKit. Apple Developer Documentation (2022) https://developer.apple.com/documentation/mapkit.
5. Taloor, A.K., Goswami, A., Bahuguna, I.M.: Remote Sensing and GIS Applications in Water Cryosphere and Climate Change. Remote Sensing Applications: Society and Environment, 28, pp. 100866 (2022) doi: 10.1016/j.rsase.2022.100866.
6. Cheng, H., Wu, S., Zheng, Y.: GIS-Based Mineral Prospectivity Mapping Using Machine Learning Methods: A Case Study from Zhuonuo Ore District, Tibet. Ore Geology Reviews, 161, pp. 105627 (2023) doi: 10.1016/j.oregeorev.2023.105627.
7. Arumugam, T., Ramachandran, S., Kinattinkara, S.: Bayesian Networks and Intelligence Technology Applied to Climate Change: An Application of Fuzzy Logic Based Simulation in Avalanche Simulation Risk Assessment Using GIS in a Western Himalayan Region. Urban Climate, 45, pp. 101272 (2022) doi: 10.1016/j.uclim.2022.101272.
8. Abdullah, S., Barua, D.: Combining Geographical Information System (GIS) and Machine Learning to Monitor and Predict Vegetation Vulnerability: An Empirical Study on Nijhum Dwip, Ecological Engineering, 178, pp. 106577 (2022) doi: 10.1016/j.ecoleng.2022.106577.
9. Sharma, S.K., Misra, S.K., Singh, J.B.: The Role of GIS-Enabled Mobile Applications in Disaster Management: A Case Analysis of Cyclone Gaja in India. International Journal of Information Management, 51, pp. 102030 (2020) doi: 10.1016/j.ijinfo-mgt.2019.10.015.
10. Gao, S., Mai, G.: Mobile GIS and Location-Based Services. Comprehensive Geographic Information Systems, pp. 384–397 (2018) doi: 10.1016/b978-0-12-409548-9.09710-4.
11. Apple: App Development with Swift. Apple Education (2018)
12. Apple: Presenta El Chip M2 Ultra. Apple Newsroom (2022) https://www.apple.com/mx/newsroom/2023/06/apple-introduces-m2-ultra/.
13. Synchronize MapView and SceneView: Sample Code: Arcgis Maps SDK for JavaScript 4.28: Arcgis developers, Sample Code ArcGIS Maps SDK for JavaScript 4.28 ArcGIS Developers (2023) https://developers.arcgis.com/javascript/latest/sample-code/views-synchro-nize/.
14. Kamoun, F., Iqbal, F., Esseghir, M.A.: AI and Machine Learning: A Mixed Blessing for Cybersecurity. In: International Symposium on Networks, Computers and Communications (ISNCC), pp. 1–7 (2020)
15. Alpaydin, E.: Machine Learning: The New AI. MIT Press (2016)

ISSN 1870-4069

Implementation of an Embedded System to Detect the Use of Face Masks Using a CNN Model for the Networking Laboratory at the UQRoo Campus Cancun

María Kristell Flores Reyes, Daniel Sergio Martínez-Ramírez,
Julio Cesar Ramírez-Pacheco, Homero Toral-Cruz,
David Ernesto Troncoso-Romero,
Ismael Osuna-Galán, J. A. León-Borges

Universidad Autónoma del Estado de Quintana Roo,
División de Ciencias, Ingeniería y Tecnología,
Mexico

jleon@uqroo.edu.mx

Abstract. The access to face to face classes at the Autonomous University of the State of Quintana Roo requires the use of a face mask detector as a personal protection measure against any risk of COVID-19 infection or other respiratory diseases in enclosed spaces. As a proposed facemask detector, the present work consists of the implementation of an embedded instrument that performs the detection of face masks inside an enclosed space such as the networking laboratory so that students and teachers have a safer space against possible health risks. Therefore, the implementation of an object detection method was sought to identify whether any student or any staff member is wearing his or her face mask to have access to the laboratory facilities. Therefore, in this research work, we implement an embedded face mask detector system based on Convolutional Neural Networks (CCN) used over a Raspberry Pi 4 computer.

Keywords: Convolutional neural networks (CNN), embedded systems, face mask detector (FMD), real time face mask detector (RTFMD), raspberry Pi 4.

1 Introduction

In the Autonomous University of the State of Quintana Roo (UQRoo) at Cancun, the use face masks were a then mandatory requirement to protect the staff and the student community from COVID-19, and to keep a reasonable distance about 1.5 metres each one including the access to the campus facilities through sanitary filters [1].

In most of most public Mexican universities, there is no a usual monitoring of the entrance facilities using a face mask detection system. In the Networking Laboratory, there is an opportunity of deploying an embedded-system prototype to detect fase masks on faces through a comparison of two face detection models based on Convolutional Neural Networks (CNN) which are analyzed in terms of metrics to choose the metric with the best performance. Once, the best choice CNN-based model is implemented on a Raspberry Pi 4 computer.

Conceptually, Artificial Intelligence (AI) consists of systemised methods that mimic the human intelligence to perform tasks and to interactively improve upon the gathered information. AI consists of a variety of fields from bionic robotics, as well as video games, its logic and data processing speed mainly achieved through artificial neural networks [2].

On the other hand, Computational Intelligence (CI) comprises the extraction, learning, reasoning, and training methods to build new knowledge. Within CI, there are two subfields which are Deep Learning (DL) and Machine Learning (ML) that supports the statistical methods to allow machines to improve experiences [3].

On the other hand, ML automatically seeks to learn meaningful relationships and patterns from examples and observations. Advances in ML have enabled the emergence of intelligent systems with human like cognitive capabilities. The ability of such systems to solve problems is based on analytical models that generate predictions, rules, answers, recommendations or similar results [4]. The following six stages summaries the process for building a ML model:

1. Collecting data: Data can be collected from sources such as a website, using either an Application Programming Interface (API) or a database. It´s important to notice that this stage is highly time consuming because data feed the chosen ML model.

2. Data Preprocessing: Once data are available, data need to be processed to ensure the adequate format to feed the ML algorithm of interest. In practice, several pre-processing tasks have to be performed before the proper use of data properly.

3. Data Exploration: This stage performs a preliminary analysis to fix any missing values cases or to try to find, at first sight, any pattern on data to make easier the model construction. At this point, outliers should be detected; or find the most influential characteristics to make a prediction.

4. Algorithm Training: Processed data feed ML algorithms with data previously processed. The aim is to extract the useful information from the initial data and then to make predictions properly.

5. Algorithm Evaluation: The evaluation of an ML algorithm generates information which tests over the information generated by the knowledge of the previous training obtained through previous interactions.

6. Algorithm Implementation: Then, the ML algorithm is finally implemented by a programe on either a Personal Computer (PC) or an embedded system.

Summarizing these stages, we find a learning paradigm, based on supervised learning that means the algorithm is taught how to perform a job, having a classified dataset under a certain appreciation or idea to find patterns that can be applied in an analysis, and produce an output that is already known. In next section, we present some key required preliminaries about Convolutional Neural Networks(CNN) to deploy this proposed present work.

2 Face Detection Systems Preliminaries

2.1. Face Detection Systems

Face detection refers to the ability to identify the presence of human faces on digital images, using ML algorithms to determine whether one or more faces without regarding "whose face is whose" to only count the number of people on the analyzed image, then these analyzed images are stored into a searchable database. Moreover, face detection algorithms often start locating human eyes, which constitute what is known as a valley region and therefore one of the easiest features to be detected. Once the eyes are detected, the algorithm might attempt to detect facial regions, including eyebrows, mouth, nose, nostrils and iris. Once the algorithm assumes that has detected a facial region. Then the algorithm, can apply additional tests to validate whether it has detected a face [6].

2.1.1. Object Detection Architectures Based on CNN Face Detection Algorithms

As CNNs are deep learning algorithms to analyze and learn visual features from large amounts of data [7]. Object detection allows computer systems to "see" their environments by detecting objects on images or visual videos. Through the use of supervised learning as seen from ML algorithms [8].

Supervised learning is when we train a ML algorithm by giving it questions (features) and answers (labels). The algorithm can make a prediction knowing the features. In this type of learning there are two training algorithms: classification and regression [8].

2.2. A Face Detection Model Using CNNs

The structure of CNNs is inspired by neurons in the human and animal brain. CNNs consists of numerous convolutional layers preceding the subsampling (clustering) layers, while the final layers are Fully-Connected (FC) [5].

Object detection allows computer systems to "see" within their views by detecting objects on either images or visual videos. This is made through the use of supervised learning as seen from ML algorithms [8].

As usual, supervised learning is when a ML algorithm is trained by avoiding the questions (features) and answers (labels). Then, the algorithm can make a prediction by knowingthe keyimage features. In this type of learning, there are two algorithms (training): Classification and regression. In [8], classification techniques expect the algorithm to say to which group under study the element of interest belongs, Sandoval mentions that the regression algorithm is a method on which a number is expected. In context, this number does not have a place in a group, but returns a specific value.

2.2.1. CNN Model Training

The classification of a supervised learning CNN-based model is performed after training the model to classify the trained images into their respective classes by

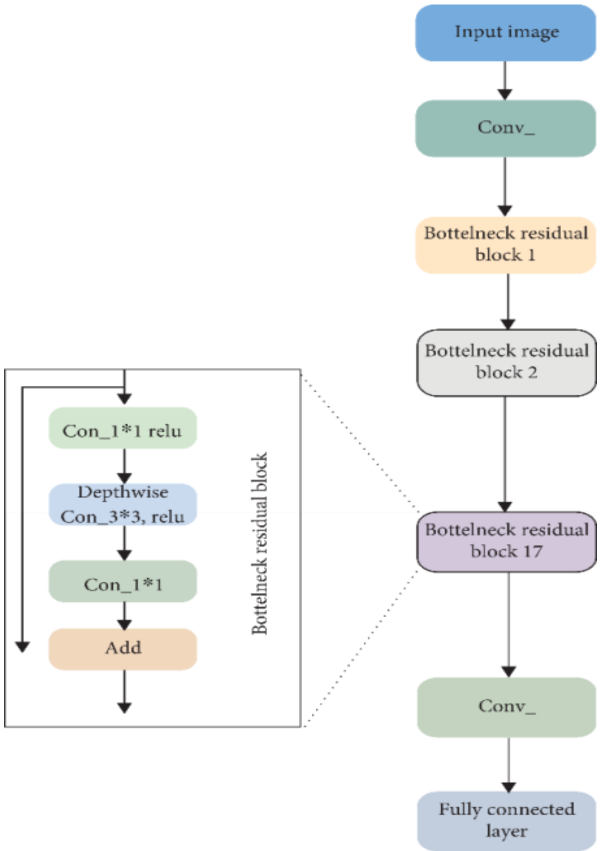

Fig. 1. MobileNet V2 architecture, reference [13-16].

learning important visual patterns. Typically, Tensorflow and Keras are the most commonly used Python-based libraries to be deployed on special computers such as embedded systems [10].

2.2.2. CNN Model for Face/Mask Detection

When the CNN-based model is trained, the model can be implemented so that any image detects the presence of a face mask as the object of interest. The gathered image is initially sent to the face detection model to detect all faces within the environment image. Then, these faces are passed as inputs to the face mask detection model. Next, the model would extract hidden patterns/features of interest from the image and finally classify them as "face masks" or "No face masks" [10].

Based on these preliminaries, we present a brief description of the architecture MobileNet V2 and used metrics, deployed on this proposed embedded system towards its definite deployment.

3 MobileNet V2 Architecture and Metrics

3.1. MobileNetV1/V2

To develop this embedded system, we choose MobilNets from Google, a CNN-based model whose core architecture is based on Depthwise Separable Convolutions including a width multiplier. For more details about MobilNets, see [12-16].

In addition, MobilNet V1/V2 are designed for mobile devices to support classification, detection. The ability to run deep networks enhances the user experience, in terms of accessibility anytime and anywhere as well as security, privacy, and energy consumption saving facilities [9].

While, MobileNetV2 is an enhancement of MobileNetV1 and serves as the state of the art for mobile visual recognition, including classification, object detection and semantic segmentation. MobileNetV2 is released as part of the Tensor Flow-Slim image classification library [9].

Moreover, we select MobileNet V2 because this architecture deals with linear bottlenecks including inverted residuals, for more details about the MobileNet V2 architecture are presented in [16-18]. In Figure 1, we present the diagram of the MobileNet V2 architecture.

3.2. Metrics of Performance

For performance of classification models, we only consider the following metrics such as precision is a non-negative prediction number that indicates how correct the system is, as seen in Equation (1), recall, or sensitivity, implies how many confident instances the model identifies in Equation (2), F1 score represents the mean of recall and precision, yielded in Equation (3) including the accuracy which highlights the usefulness and reliability of the used method as presented in Equation (4).

They are presented from Equations (1) to (4) respectively:

$$Precision = \frac{TP}{FP+TP}, \tag{1}$$

$$Recall = \frac{TP}{FN+TP}, \tag{2}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{3}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}. \tag{4}$$

where FN, FP, and TP represent the False Negative, False Positive, and True Positive samples respectively, from an internal confusion matrix as originally presented in [13-15].

For this embedded system, we make use of the of the Face Mask Detector (FMD) method and the Real Time Face Mask Detector (RTFMD) method besides a performance comparative of metrics in terms of the previously mentioned metrics.

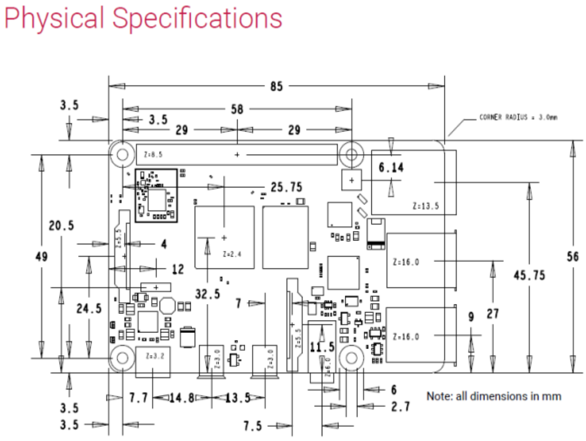

Fig. 2. Diagram of the physical specifications of the **Raspberry Pi 4** computer.

Fig. 3. The Raspberry Pi 4 model B computer.

4 Implementation and Technical Challenges

4.1. The Raspberry Pi 4 Computer Platform

For the implementation of this proposed embedded system, we require a platform to deploy the face detection models using the FDM metrics for CCNs. As the first technical challenge, we choose a portable computer which can load an mobile Operating System (OS) and we decide to use an open-source OS, a community developed one, such that the Raspbian OS, a light version based on Linux Debian OS, as the basis for the installation of the corresponding Python libraries to evaluate the FDM metrics.

As second technical challenge, we selected the Raspberry Pi 4 computer because this type of computer provides the possibility to connect input/output hardware such as monitors including some outstanding built in features such as USB connectivity,

Table 1. Technical specifications of the Raspberry Pi 4 Model B computer, see [11].

Specification	
Processor:	Broadcom BCM2711, quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz
Memory:	1GB, 2GB, 4GB or 8GB LPDDR4 (depending on model) with on-die ECC
Connectivity:	2.4 GHz and 5.0 GHz IEEE 802.11b/g/n/ac wireless LAN, Bluetooth 5.0, BLE Gigabit Ethernet 2 × USB 3.0 ports 2 × USB 2.0 ports.
GPIO:	Standard 40-pin GPIO header (Fully backwards compatible with previous boards)
Video & sound:	2 × micro-HDMI ports (up to 4Kp60 supported) 2-lane MIPI DSI display port 2-lane MIPI CSI camera port 4-pole stereo audio and composite video port
Multimedia:	H.265 (4Kp60 decode); H.264 (1080p60 decode, 1080p30 en-code); OpenGL ES, 3.0 graphics
SD card support:	32 GB Micro SD card slot for loading operating system and data storage
Input power:	5V DC via USB-C connector (minimum 3A) 5V DC via GPIO header (minimum 3A) Power over Ethernet (PoE)–enabled (Requires separate PoE HAT)
Environment:	Operating temperature 0–50ºC
Compliance:	For a full list of local and regional prod-uct approvals, please visit https://www.raspberrypi.org/documeta-tion/hardware/raspberrypi/conform-ity.md
Production lifetime:	The Raspberry Pi 4 Model B will remain in production until at least January 2026

Gigabit Ethernet including Wireless LAN connectivity. In Figure 2, we yield the physical specifications of the Raspberry Pi 4 computer [11].

Based on its physical specifications, this computer was selected because of its credit card size, providing to our embedded system a compact size to provide an easier installation. According to the Raspberry site, the model used in this implementation is the Raspberry Pi 4 Model B. An illustration of the Model B is shown in Figure 3 below.

Next, we present the full technical specifications of the Raspberry Pi 4 Model B computer below. Based in Table 1, the model B was selected because of its specific features, RAM capacity, processor, and SD card support which provides the storage to install the OS to be loaded in the system.

4.2. CNN-Based Face Mask Detector Method

The implementation of CNNs using the FMD-based method starts with the installation of the operating system that is obtained on the Raspberry Pi Imager,

Table 2. Installed software dependencies for the FMD methods [18].

Dependencies	Name of the architecture or version
System architecture	Aarch64
Python	3.9.2
TensorFlow	2.11.0
Keras	2.11.0
Time-Python	0.0.15
OpenCv-Python	3.1.0
Numpy	1.23.5
Picamera2	0.3.6
Imutils	0.5.4

which can be downloadeddirectlyfromtheRaspberry web site. Afterdownloading, theuserexecutes the Raspberry Pi image installation file on the systems in order to install the OS. Then, the OS is ready to be implemented on the 32 GB microSD memory. Then, the Raspberry Pi OS (64-bit) Desktop (compatible with Raspberry Pi 4) is installed.

Once the operating system has been installed on the Raspberry Pi 4 microSD memory and the OS starts to be configured. Once the Graphic User Interface (GUI) started, the installation of the dependencies for the FDM method starts installing on the system.

Installation of the FMD Software Dependencies

Once the FDM is running, the user installs the main software dependencies to operate the FDM method from the terminal of the Raspberry Pi OS (Raspbian). Table 2 presents the required key software dependencies to be installed and configured on our embedded system [18].

4.3. Installation and Activation of the Raspberry Pi 4 Camera

To activate the Raspberry Pi camera, the user types the following command: sudo raspiconfig, on the terminal. Once the camera is activated, the installation process starts takes place. The next configuration windows appear as shown in Figure 4a), by selecting option 5 which shows of the interface options, where the peripherals can be configured.

When the Interfacing Option menu is chosen, the user accesses to either activating or deactivating the menu choices. Next, Figure 4b) shows the first option to activate the Raspberry Pi camera connection.

Once the Interfacing Option is selected. Subsequently 5c) shows the first option to activate the connection to the Raspberry Pi camera. Finally, Figure 5d) indicates the option to reboot the system to save the configuration data, to make effect configuration changes.

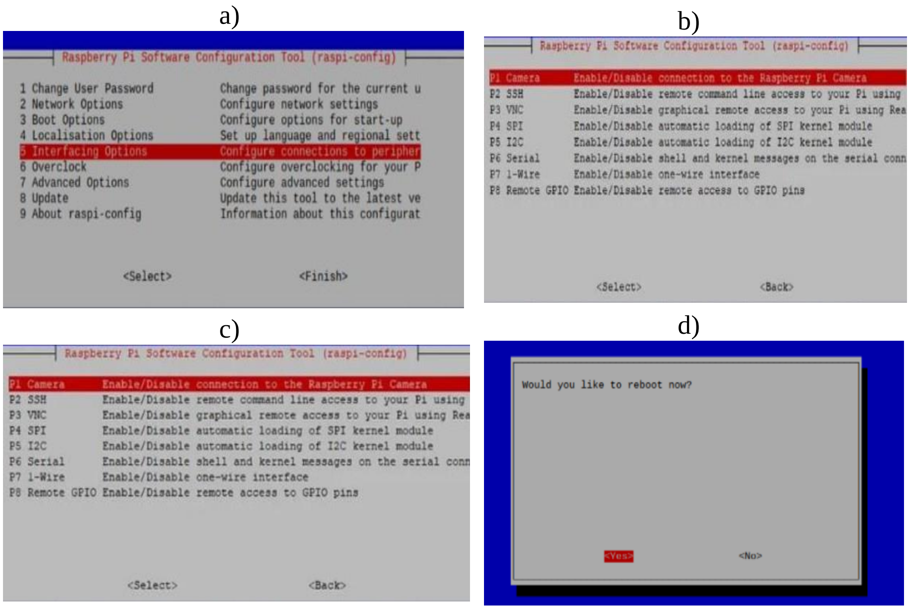

Fig. 4. a) The Raspberry Pi software configuration tool interface. b) The Raspberry Pi Software Configuration Tool menu.c) The Raspberry Pi Software Configuration Tool menu, the user selects the "P1 Camera" option and then choosing the enable option. d) The system reboot window to activate the camera [18].

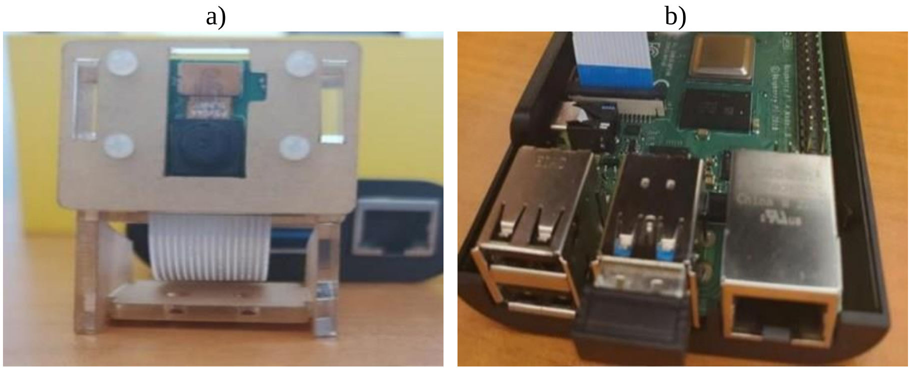

Fig. 5. a) The used Raspberry pi camera. b) The connection between the camera and the Raspberry Pi computer: The camera cable is plugged into the internal port [18].

4.4. Implementation of the Raspberry Pi

Figure 5a) yields the built-up camera for the Raspberry Pi 4 computer while Figure 5b) shows the connection of the flexible cable from the camera to the Raspberry Pi 4 Model B [18].

a)

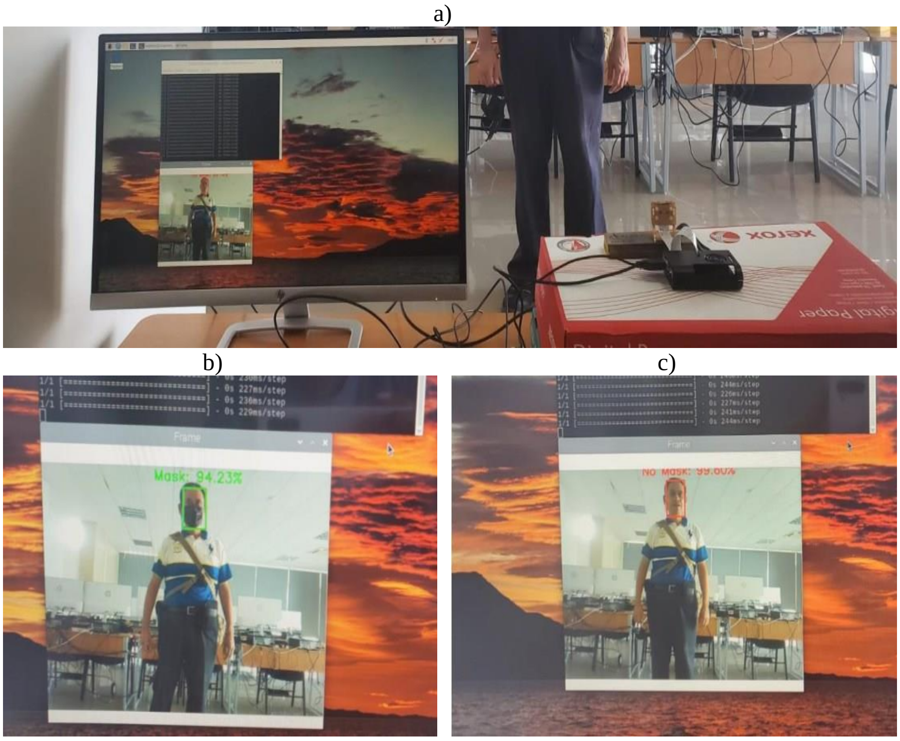

Fig. 6. a) The peripheral connection deployment: monitor with a micro-HDMI cable, a keyboard, a mouse, and a supply cable. b) Face detection using a face mask. c) Face detection without a face mask [18].

As physical implementation, Figure 6a) shows the Raspberry Pi 4 deployment at the Networking Lab. In the same Figure 6a), the monitor shows the output image and a green box appear for wearing a face mask.

In this deployment [18], the user opens the terminal over the directory of the CNN model "Face Mask Detector", made by the following command line; cd face-mask-detector-main. Subsequently, the main Python script is executed to start the face mask detection system, with the following python interface command; python3 detect_mask_video.py

After a few seconds, the camera frame opens to detect whether or not the face mask is present. Figure 6b) presents a zoom-in of the monitor which shows the detection of face mask on face in real time by a green-colored box with the state description "Mask".

Figure 6c) indicates the face detection without a face mask, showing a red-colored box with the description "No Mask". Figures 7b) and 7c) yield how the detection of the use of face masks on real time.

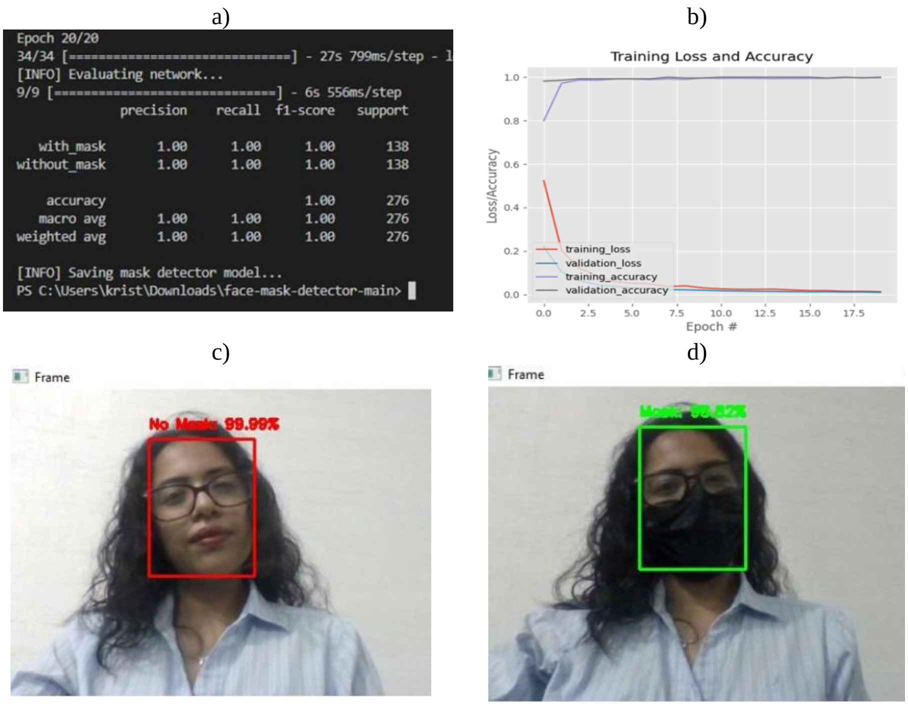

Fig. 7. Metrics of the *Face Mask Detector* **FMD** code. a) Performance metrics b) Graph of losses and accuracy for the **FMD** code. c) Face detection without face mask, **FMD** model. d) Face detection with face mask [18].

5 Results and Discussion

5.1. Results

5.1.1. Method 1: Face Mask Detection (FMD)

Based on the image dataset experiments in [5] which presents both conditions: with mask and without mask, as “with_mask” and “without_mask” respectively. We present the FMD model to obtain the previously mentioned performance metrics: recall, precision, and F1-score. Running the training model, showed in Figure 7, we yield the performance metrics of results [18].

Figure 7a) shows the classification of performance metrics for precision, recall, and F1-score by, sorted by columns, respectively while both with/without face mask conditions including the accuracy are yielded as rows up to 20 epochs with a running time of about 7 seconds. Figure 7b) presents the training/validation accuracy in terms of the number of epochs which tends to the unity from 2.5 epochs while the

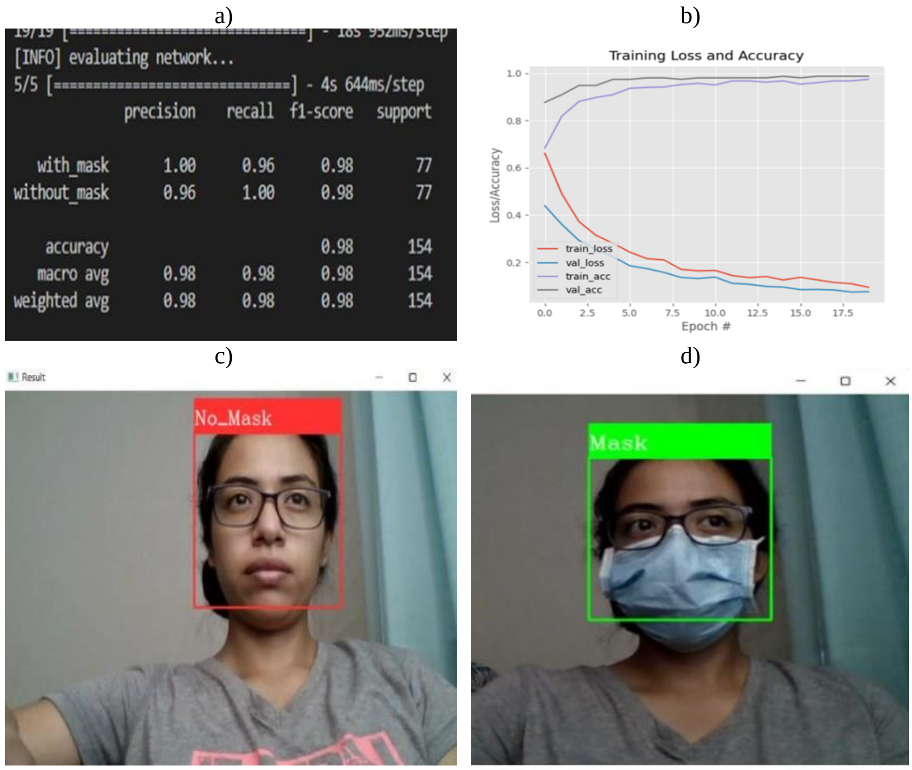

Fig. 8. Metrics of the *Face Mask Detector* **RTFMD** code. a) Performance metrics b) Graph of losses and accuracy for the **RTFMD** code. c) Face detection without face mask, **RTFMD** model. d) Face detection with face mask [18].

training/validation loss has a decreasing behavior. Meanwhile, Figures 1c) and 1d) show the results of the FMD method without and with face masks as detection results: a red frame for the "No Mask" status and a green frame "Mask", respectively [18].

5.1.2. Method 2: Real Time Face Mask Detection (RTFDM) Training

On the other hand, the Real Time Face Mask Detection RTFMD method project does not contain a dataset. However, the dataset was downloaded from [17], and attached to Naemazam's project (2022) [17]. Based on [19], we prepared a dataset having two folders named "with_mask" and "without_mask" image databases. By running the training model, as viewed in Figure 8, presents the measurement of metrics.

In this point, we present the RTFMD method to obtain the previously mentioned performance metrics: recall, precision, and F1-score. Running the training model, showed in Figure 8, we yield the performance metrics of results [18].

Analogously with the previous subsection, Figure 8a) shows the classification of performance metrics for precision, recall, and F1-score by, sorted by columns, respectively while both with/without face mask conditions including the accuracy are yielded as rows up to 20 epochs with a running time of about 5 seconds. Figure 7b)

Table 3. A comparison of metrics [18].

Name of the model	With or without face masks	Precision	Recall	F1-Score
Model 1 FMD	With face masks	1	1	1
	With no face masks	1	1	1
Model 2 RTFMD	With face maks	1	.96	.98
Accuracy = 0.98	With no face masks	.96	1	.98

presents the training/validation accuracy in terms of the number of epochs which carries out behaviors less than the unity while the training/validation loss have a decreasing be-haviour. Meanwhile, Figures 1c) and 1d) show the results of the RTFMD method with-out and with face masks as detection results: a red frame for the "No Mask" status and a green frame "Mask", respectively [18].

5.2. A Comparison of Metrics of the Face Mask Detector (FMD) and the Real Time Face Mask (RTFMD) Detection methods

5.2.1. Discussions

For discussions of results, we make a performance comparison between the FMD and the RTFMD models to choose the model to be deployed on the Raspberry Pi 4 computer. Table 3 yields this comparison in terms of precision, recall, and F1-score performance metrics for both with and without detection of face masks, tested on both academic staff and student members.

In addition, Table 3shows the precision, recall, andF1-scoremetrics for both the FMD and the RTFMD methods which cover performance metrics in a range from 90% to 100% [18].

Based on the results presented in Table 3, section four deals with the implementation of the FMD model on our home-made face detection embedded system using the Raspberry Pi 4 Model B.

Based on performance results, the FDM method was chosen because it showed the best accuracy in detecting face masks closely to the 100% even though the RTFDM method showed an accuracy of about 98%. It is worth mentioning that, although this embedded instrument is under its experimental deployment [18].

6 Conclusions

The implementation of an embedded face detection system using mouth covers based on a Raspberry Pi 4 platform allows for a reliable, programmable, expandable and upgradeable instrument that can implement face detection algorithms based on CNNs.

Then, it is possible to evaluate different face detection modelsover the same instrument to help preventing of infectious diseases not only the covid-19 but also to help implementing protection and prevention measures within a crowded space.

As based on Raspberry Pi 4, we can evaluate not only itsconfiguration and installation of input/output devices managed by a Linux-based OS and implemented in Python-based libraries to be able to perform and evaluate different face detection models.

As a practical evaluation, this implementation of the face cover detection system over a Raspberry Pi 4 computer. However, there is not still a storage method to keep image databases because of the storage limitations of this small computer. Nowadays, there is an ethical concern of image uses but the purpose of this embedded system is only to detect face masks on faces. For deployment places, it is originally implemented in the networking laboratory but this embedded computer would be applied in other parts of campus.

Viewing the RTFDM performance, the accuracy was less effective than the FDM model. However, this embedded system would use different detection models or better. For detection optics, this embedded system does not originally consider camera optics parameters such as lightning, viewing angles which would affect the performance in case of groups of students arriving to the laboratory. For design criteria, this embedded system is conceived to operate indoors instead of outdoor because we assumed lightning conditions as constant. To consider angle views, these views would depend on the camera location.

Although the Raspberry Pi 4 architecture allows connectivity to the Internet, the focus of this work is the comparison of face detection metrics using FDM and RTFDM algorithms based on CNNs using MobileNet V2. The scope of this instrument design is based on a functional embedded system that can be implemented at a very reasonable cost for applications in enclosed, confined and crowded spaces. Finally, we should mention that this face detection system is not currently an IoT device at this moment, but thanks to the versatility of the Raspberry Pi 4 platform, we can add the IoT functionality in future to ensure a cloud storage of images through a local campus server.

References

1. Universidad Autónoma de Quintana Roo: Protocolo y Plan de Seguridad Sanitaria (2020) https://www.uqroo.mx/imagen2020/COVID%2019/PROTOCOLO_DE_SEGURIDAD_SANITARIA-UQROO.pdf.
2. Corporación Universitaria Republicana: Artificial Intelligence at a Glance (2022) http://ojs.ure-publicana.edu.co/index.php/ingenieria/article/view/234/213.
3. Gonzalez, L.: Inteligencia Artificial vs Machine Learning vs Deep Learning. Learn AI (2021) https://aprendeia.com/inteligencia-artificial-vs-machine-learn-ing-vs-deep-learning.
4. Russell, S., Norvig, P.: Artificial Intelligence: A modern approach. Pearson Series In Artificial Intelligence (2021)
5. Alzubaidi, L., Zhang, J., Humaidi, A.J.: Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions, a Survey Paper. Journal of Big Data (2021) doi: 10.1186/s40537-021-00444-8.

6. Diferencia entre face detection y face recognition: Tvc. MX (2020) http://soporte.tvc.mx/Ingenieria/DAHUA/ARCHIVOS_COMUNES/Diferencia%20entre%20Face%20Detecti%C3%B3n%20y%20Face%20Recogniti%C3%B3n%20Dahua.pdf.
7. What Is Computer Vision?: Intel (2023) https://www.intel.com/content/www/us/en/internet-of-things/computer-vision/convolutional-neural-networks.html.
8. Sandoval, L.J.: Machine Learning Algorithms for Data Analysis and Prediction ITCA-FEPADE (2018) http://www.redicces.org.sv/jspui/bitstream/10972/3626/1/Art6_RT2018.pdf.
9. MobileNetV2: The Next Generation of on-Device Computer Vision Networks. Google-blog.com (2020) https://research.google/blog/mobilenetv2-the-next-generation-of-on-device-computer-vision-networks/.
10. Goyal, H., Sidana, K., Singh, C.: A Real Time Face Mask Detection System Using Convolutional Neural Network. Multimedia Tools and Applications, 81(11), pp. 14999–15015 (2022) doi: 10.1007/s11042-022-12166-x.
11. Model, B.: Raspberry Pi 4 Computer (2023) https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-product-brief.pdf.
12. Howard, A.G., Zhu, M., Chen, B.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications (2017) doi: 10.48550/arXiv.1704.04861.
13. Hussain, D., Ismail, M., Hussain, I.: Face Mask Detection Using Deep Convolutional Neural Network and MobileNetV2-Based Transfer Learning. Wireless Communications and Mobile Computing, pp. 1–10 (2022) doi: 10.1155/2022/1536318.
14. Asghar, M. Z., Albogamy, F.R., Al-Rakhami, M.S.: Facial Mask Detection Using Depthwise Separable Convolutional Neural Network Model During COVID-19 Pandemic. Frontiers in Public Health, 10 (2022) doi: 10.3389/fpubh.2022.855254.
15. Farman, H., Khan, T., Khan, Z.: Real-Time Face Mask Detection to Ensure COVID-19 Precautionary Measures in the Developing Countries. Applied Sciences, (12)8, pp. 3879, (2022) doi: 10.3390/app12083879.
16. Sandler, M., Howard, A., Zhu, M.: MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition, (2018) doi: 10.1109/cvpr.2018.00474.
17. Solano, G.: Data set fase (2021) https://github.com/GabySol/OmesTutorials2021/tree/main/Mascarillas%20 dataset/Dataset_faces/Sin_mascarilla.
18. Flores-Reyes, M.K.: Implementación de un sistema de detección del uso de cubrebocas usando un modelo CNN en una Raspberry Pi para el Laboratorio de Redes en la UAEQRoo Campus Cancún (2022)
19. Azam, N.: Real Time Face Mask Detection (2022) https://github.com/naemazam/Real-Time-Face-Mask-Detection.

Electronic edition
Available online: http://www.rcs.cic.ipn.mx

http://rcs.cic.ipn.mx

Centro de Investigación
en Computación