Data Science for Business - Group Assignment

1. Introduction

1.1 Application name:

(62 characters)

Optical Character Recognition system for billet identification

1.2 Abstract:

(307 characters)

This project aims to improve the billet identification process used at SMS Concast. Recent improvements in hardware devices and in machine learning models open the potential to improve the logistic tracking system. This could result in potential cost savings, market expansion and reducing CO2 emissions.

1.3 Executive Summary:

(1456 characters)

SMS Concast, a trailblazer in the steel manufacturing industry, is embarking on a significant digital transformation initiative. This endeavor is not just a technological upgrade but a strategic move to redefine the operational landscape of steel manufacturing. At the core of this initiative is a focus on leveraging digital technologies to overcome longstanding challenges in the industry, including inefficient identification processes and the overarching need for more sustainable manufacturing practices.

The project represents a monumental shift from traditional methodologies to a data-driven, technologically advanced approach. A key component of this transformation is the intensive data collection and labeling process. This foundational step involves gathering extensive datasets, encompassing detailed imagery and production metrics. The data, meticulously collected and labeled, will serve as the groundwork for training sophisticated machine learning models. These models are pivotal to the initiative, as they will drive the advanced identification systems set to be deployed.

The goal of this digital transformation is multifaceted. Primarily, it aims to significantly enhance the accuracy and efficiency of steel billet identification processes, a critical aspect of steel manufacturing that has been mired in challenges. By doing so, the initiative promises not only to boost productivity but also to reduce operational costs significantly.

Part 2. Organizations and People

(6748 characters)

2.1 Research Partners

Contact Party	Research Center Type	Department	Organisation Representative	Contact
CSEM - Centre Suisse d’éléctronique et de microtechnique	Research facilities of national importance (FNB)	Edge AI & Vision	Andrea Dunbar	Jaquet-Droz 1 2000 neuchatel andrea.dunbar@csem.ch +41 78 911 53 21 +41 32 720 50 69

Contact Party	Department	Organisation Representative	Contact
SMS Concast AG	Caster concepts&Technological&Digital solutions	SMS group management	Tödistrasse 9 8027 Zürich Switzerland +41 44 204 65 11 Mconcast.info(at)sms-group.com

2.2 Implementation Partners2.3 Roles and Responsibilities

Roles	Responsibilities	Source
Project Manager	• Responsible for overall project planning, team coordination, resource allocation, and progress tracking	CSEM
Research Partner Representative	• Ensures that research activities align with project objectives, providing expertise and technical support	CSEM

Technical Team	• Guides the technical team in developing solutions, ensuring the implementation meets project requirements	CSEM,SMS Concast AG
Head of supply chain	• Manages and optimizes a company's supply chain and logistics, overseeing strategy and operations to enhance efficiency and ensure smooth delivery of products or services.	SMS Concast AG
Market and Business Strategy Consultant	• Analyzes market trends and competitive environment, formulating effective market entry and commercialization strategies.	SMS Concast AG
Risk Management Consultant	• Identifies potential risks, develops and implements risk management plans	CSEM
Data Analysis Expert	• Oversees the product development process to ensure all products meet quality and safety standards	CSEM

Part 3 Value Creation

In the steel industry, the tracking system is extremely important, as one fasle quality steel bullet can lead to a huge problem, such as needing to recall around 1000 cars or other final products and harming the company's reputation. Moreover, it will generate CO2 when recalled and will be wasted in the steel milling system. Every year, 150,000 billets will be melted in the milling industry, which is 270000 tons of steel, and 2% of this still needs to be remelted at the cost of 725kWH and the cost of 83 CHF due to identification error. It faces the imperative of digital transformation to maintain competitiveness and efficiency. Among the key players in this transition is SMS Concast AG, a company at the forefront of adopting digital strategies. In collaboration with CSEM, SMS Concast AG embarks on a groundbreaking initiative to revolutionize its operations, starting with an enhanced tracking system. This essay aims to dissect the nuances of this digital transformation, scrutinizing the business model of SMS Concast AG, particularly in its collaboration with CSEM, and its broader impact on the steel industry.

3.1 Business Targets and Model

SMS Concast AG's digital transformation initiative, powered by CSEM's expertise, pivots on several business objectives. Foremost is the enhancement of operational efficiency. The integration of advanced tracking systems aims to streamline production processes, significantly reducing time and resource wastage. This efficiency drive is coupled with a strong focus on cost reduction, leveraging technology to minimize operational expenses and improve bottom-line performance.

In terms of business modeling, SMS Concast AG seeks to position itself not just as a steel manufacturer but as a technology-driven solutions provider. This shift entails a strategic move up the value chain, offering clients not only steel products but also technological expertise. The partnership with CSEM, renowned for its prowess in OCR and AI technologies, provides a cutting-edge dimension to the company's offerings. These technological innovations are at the heart of SMS Concast AG's new business model, promising to elevate its market standing and redefine client relationships.

3.2 Market Dynamics and Competitive Positioning

The steel industry's market dynamics present both challenges and opportunities for SMS Concast AG. As the sector increasingly embraces digital solutions, the market size for technologically advanced steel manufacturing processes grows, offering vast potential for SMS's new tracking system. This system, characterized by its unique integration of OCR and AI, positions SMS Concast AG competitively in the market, distinguishing it from traditional players. Its unique selling proposition lies in its ability to provide unparalleled efficiency and accuracy with 99.98% in steel production.

However, the path is not without challenges. SMS Concast AG must navigate a market replete with established players and emerging competitors, all vying for technological supremacy. The company's ability to leverage its USP and effectively communicate its benefits to potential clients will be crucial in maintaining its competitive edge.

3.3 Economic Value

The economic value of SMS Concast AG's digital initiative is profound. With the implementation of CSEM's advanced tracking system, the project's Net Present Value (NPV) is expected to be significantly positive, indicating a lucrative return on investment. This financial benefit is complemented by cost savings of 533,925 CHF from improved operational efficiencies and reduced waste.We know that every year, the milling industry metals 150,000 billets, creating 270000 tons of steel, and 2% of this still needs to be remelted at the cost of 83 CHF and 725kWH due to the error identification rate. **We assure that the remelting rate from 2% decreases to 0.067%**, So we can calculate the following：

Remelting Cost:270000 tons*1.933%*83 CHF = 433,185CHF

Electricity Used:270000 tons*1.933%*725kWH = 3783,8 MWH

We assume that the cost of electricity is 131 MWH:

Electricity Cost: 3783.8*131 = 495,677 CHF

Total Cost: 433,185+495,677= 92,8862 CHF

3.4 Customer Model and Market Access Strategy

SMS Concast AG's approach to market access in its digital transformation hinges on a B2B customer model. The focus is on engaging with other steel manufacturing companies that would benefit significantly from the advanced tracking system. To effectively reach these clients, the company employs a multi-faceted marketing strategy that includes digital marketing, participation in industry events, and leveraging existing business relationships. This strategy is complemented by a robust implementation plan, which involves phased deployment and integration of the tracking system, ensuring smooth adoption by clients. SMS's approach is aimed at establishing long-term partnerships, emphasizing the added value and efficiency gains their clients would achieve through this technological upgrade.

3.5 Social and Ecological Relevance

The digital transformation of SMS Concast AG, particularly through its collaboration with CSEM, holds significant social and ecological implications. The project aims for economic gains and contributes to environmental sustainability in the steel manufacturing industry. If we apply this product every year, we can reduce the electricity by 37,83.8 MKH. We know that Electricity produced in Switzerland has a carbon intensity of 33 gCO2/kWh. So totally, it will reduce 124.8654 ton CO2. the The advanced tracking system reduces resource wastage and energy consumption, aligning with global efforts to combat climate change. Socially, the initiative sets a precedent in the industry, promoting technological advancement and sustainable practices. By showcasing responsible production methods, SMS Concast AG positions itself as a leader in ecological stewardship within the industrial sector.

3.6 Partnership Dynamics and Future Prospects

The partnership between SMS Concast AG and CSEM represents a synergistic collaboration, combining SMS's expertise in steel manufacturing with CSEM's technological prowess. This partnership is pivotal in driving the project's success and setting new industry benchmarks. Looking ahead, the potential of this initiative extends beyond immediate financial and operational improvements. It signifies a transformation in the industry, steering towards a more digital, efficient, and sustainable future. As this project unfolds, SMS Concast AG is poised not only to benefit internally but also to influence the broader steel manufacturing sector, driving innovation and ecological responsibility.

This essay has explored the various facets of SMS Concast AG's digital transformation initiative, underscoring its significance in the steel industry's evolution. By embracing technological advancements and sustainable practices, SMS Concast AG is charting a course for a future that balances economic success with social and environmental consciousness.

4. Solution

(15063 characters)

4.1 Describe the state of the art that you intend to advance:

In the realm of steel manufacturing, SMS Concast exemplifies a progressive merger of traditional practices with next-generation digital technologies, a synergy that stands to redefine the industry's benchmarks for productivity and quality. The advancements in continuous casting technology championed by SMS Concast have yielded significant benefits, notably in reducing operational costs while simultaneously elevating the quality and diversity of steel products available on the market.

Central to this transformation has been the incorporation of digitalization initiatives, particularly in the realm of quality control. By harnessing Optical Character Recognition (OCR) and pioneering surface recognition algorithms, the company has sharpened its ability to identify and classify billets with greater accuracy, a move that enhances the entire production chain's efficiency. The integration of Optical Product Recognition (OPR) with a sophisticated optical product measuring system represents a leap forward in precision and process control.

However, the journey is not without its challenges. The complexities of logistics within the steel manufacturing domain necessitate innovative solutions to streamline tracking, mitigate the risk of product recalls, and minimize wastage—efforts that are intricately tied to the industry's broader commitment to reducing its carbon footprint.

The potential integration of Artificial Intelligence (AI) and machine learning stands as a testament to the industry's forward-thinking ethos. These technologies, which have already revolutionized fields such as medical diagnostics, are poised to bring a similar transformative impact to steel manufacturing. The ability to process vast datasets and refine production processes could usher in an era of unprecedented operational precision and efficiency.

Looking to the horizon, SMS Concast is not content to rest on its laurels. The company is actively seeking out partnerships and is open to integrating more sophisticated AI-driven systems for real-time tracking and quality assessment. These advancements promise to catalyze a revolution in the steel industry akin to the transformative changes seen in the medical field through real-time AI-MSI systems.

Ultimately, SMS Concast's ambitions crystallize around a singular vision: to expand the horizons of steel manufacturing through a steadfast commitment to digitalization, advanced process control, and the integration of cutting-edge technologies. This vision is not merely about maintaining quality and efficiency; it's about setting a new standard for sustainability and innovation in an industry that is foundational to the global economy.

4.2 Describe the novelty of your solution (technology, product, business model or process, service):

The novelty of the solution provided by SMS Concast in the field of steel manufacturing lies in its integration of advanced digital technologies into traditional steelmaking processes.
Improving on the existing CSEM solution involves developing a more cost-effective hardware system.
Incorporating **Raspberry Pi technology** into the enhancement of the existing CSEM solution brings a new dimension of cost-effectiveness and versatility to the hardware system. The Raspberry Pi (see *Figure 1.*), known for its compact size and powerful processing capabilities, can be utilized to control **a single integrated camera system equipped with software-controlled RGB LED flashes**. This integration significantly simplifies the hardware setup, as it combines multiple functionalities into a single, efficient unit, reducing the need for separate components and thereby lowering overall costs.

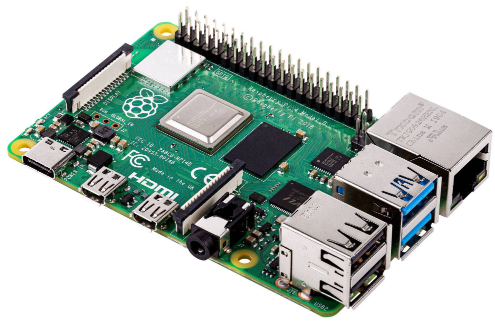

Figure 1. Raspberry Pi

The Raspberry Pi's camera module, particularly suited for this application, offers high-resolution imaging capabilities essential for accurate billet identification. This camera, coupled with the Raspberry Pi's processing power, is ideal for handling the machine learning algorithms that are integral to the system. These algorithms are enhanced to focus on fewer, yet more relevant, spectral lines, ensuring precision in identification while minimizing data processing requirements.

The application of Raspberry Pi technology also offers the flexibility to incorporate advanced features like Optical Character Recognition (OCR) using open-source tools like Tesseract. This can further refine the identification process, enabling the system to read and interpret alphanumeric data on the billets with high accuracy. Additionally, the Raspberry Pi's capability to support various programming environments and interfaces allows for easy customization and integration into the manufacturing process.

Visualizing this new hardware system, one can imagine a compact camera unit (see *Figure 2.*), powered and controlled by a Raspberry Pi, with RGB LEDs strategically arranged around the lens . This arrangement provides optimal illumination for the billet identification process. The unit's small size ensures it can be flexibly deployed throughout the manufacturing process, adapting to various operational requirements without compromising on performance or efficiency.

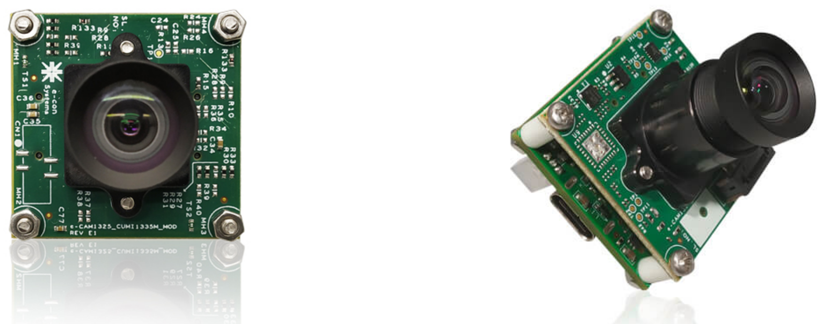

Figure 2. RGB camera solution for Raspberry Pi

By borrowing concepts from medical diagnostic technologies, such as multispectral imaging (MSI) and artificial intelligence (AI), the solution applies a **cross-disciplinary approach**.

This cross-field innovation adapts complex medical imaging techniques for use in steel manufacturing, a novel application within the industry.

Compared to the international landscape, SMS Concast aims to set a new standard for digital transformation in steel manufacturing. The international state of the art may focus on incremental improvements, whereas SMS Concast is pushing for a more radical digital leap.

The technological ambition is high as it involves integrating AI and machine learning into the production process, which could be classified at a Technology Readiness Level (TRL) of 4-6, where the technology has been validated in a lab environment but not yet in an operational setting. The risk is also significant, given the complex nature of steel manufacturing processes but also the difference between the lab and real-world conditions.

The immediate applicability is in improving the efficiency and accuracy of steel billet production, but the technology also has implications for inventory management, logistics, and supply chain optimization.

Beyond SMS Concast, the solution has potential applicability across the steel manufacturing industry and could set a precedent for digital and AI integration into other areas of heavy industry. The particular use of Raspberry Pi opens up possibilities for creative integrations. For example, the OCR output can be transformed into audio through text-to-speech (TTS) technology, broadening the scope of applications and making the system more adaptable to different industrial environments.

4.3 Describe the quantifiable goals to reach:

When setting quantifiable goals for a transformative solution like the one proposed for SMS Concast, it's essential to focus on measurable outcomes across scientific, economic, societal, and technological dimensions.

The scientific objective of this initiative is to elevate the accuracy of billet identification to a level that is virtually infallible, striving for an ambitious 99.8% accuracy rate. This leap forward in precision aims to virtually eliminate the misidentification issues that currently lead to material waste and processing inefficiencies.

From an economic perspective, the deployment of this technology is expected to translate into a direct reduction of operational costs associated with the identification process. The target is a 50% reduction in the time required for billet identification, which is anticipated to result in a corresponding increase in the throughput of the manufacturing process. By streamlining the identification phase, the overall productivity of the facility is projected to enhance by at least 30%.

In terms of societal impact, the advanced system is expected to contribute to environmental sustainability goals by reducing waste and CO2 emissions generated from incorrect billet processing by a target of 25%. Note that one false quality steel bullet can lead to the recall of several 1000 cars and the loss of reputation of their client’s company. Moreover, by automating the identification process, the goal is to improve workplace safety by reducing human interaction with heavy machinery by up to 50%, thus decreasing the likelihood of workplace accidents.

Technologically, the system sets out to achieve real-time processing capabilities with image acquisition and decision-making completed in under 3 seconds per billet. The technological integration goal is to ensure compatibility with 100% of existing production lines, allowing for a seamless transition to the new system without interrupting manufacturing operations. With regards to the machine learning component, the ambition is to train the model to an accuracy > 99.8% using a dataset of at least 100 labeled images, which would be a substantial repository given the specificity of the application.

4.4 Describe the preliminary work already performed:

The journey of SMS Concast's digital transformation has been paved by a series of foundational activities that are setting the stage for a radical upgrade in steel manufacturing processes. At the core of their preliminary work lies robust research and development. The initial stages of R&D have been focused on assessing the feasibility of incorporating Optical Character Recognition (OCR) and surface recognition algorithms within the manufacturing framework. This phase typically includes small-scale prototyping and the initial development of algorithms, serving as the proving ground for larger-scale applications.

A crucial aspect of this preparatory phase is the data collection effort. SMS Concast has been diligently gathering extensive datasets, which include high-resolution billet images and detailed production data. This rich repository of information is indispensable for the training of sophisticated machine learning models that will drive the new identification systems.

Recognizing the complexities of this technological leap, SMS Concast has actively sought out and established strategic partnerships. These collaborations with leading tech firms and academic research institutions are designed to harness external expertise in artificial intelligence and machine learning, ensuring that the solutions developed are at the forefront of technological innovation.

Pilot studies have also been a key component of their approach. These small-scale implementations serve as a critical testing ground, allowing SMS Concast to gauge the effectiveness of the new technologies within their specific manufacturing environment. It's an opportunity to iterate and refine before a broader roll-out.

Parallel to these technological endeavors, SMS Concast has not overlooked the human element. Preliminary training programs for staff have been introduced, aimed at smoothing the transition to new digital tools and workflows. This ensures that the workforce is not only comfortable but also proficient with the emerging technologies that will soon be part of their everyday operations.

Underpinning all these advancements is the necessary upgrade of the company's technology infrastructure. The IT systems are being bolstered to handle the substantial increase in data processing and connectivity that these new technologies demand.

Moreover, regulatory compliance has been a guiding principle throughout this process. Every new technological implementation is being rigorously assessed to ensure it meets the stringent standards and regulations that govern the industry.

Lastly, with an eye on sustainability, SMS Concast has conducted initial environmental assessments to gauge the impact of their digital transformation. These evaluations are crucial in developing strategies that not only enhance efficiency but also contribute to the company's environmental responsibility goals.

These preliminary activities are not merely steps but are in fact critical building blocks for SMS Concast's ambitious digital transformation. They create a solid foundation from which the company can launch its full-scale digital strategy, addressing any potential challenges early on and setting the stage for a seamless transition into a more technologically advanced manufacturing era.

4.5 Describe how the selected partners are suited for the planned project:

When selecting partners for a transformative project such as the one SMS Concast is embarking on, it's crucial to evaluate a range of competencies and synergies that these collaborators can bring to the table. For a project rooted in Switzerland, the chosen partners must not only have an established operational presence within the country but also be actively involved in enriching the local economy and the broader innovation ecosystem. Their experience in working within the Swiss industrial sector and navigating the local market nuances and regulatory frameworks will be invaluable, ensuring that the project's ambitions align with national economic and innovation strategies.

Furthermore, the scientific and technical prowess of the partners is paramount. Their expertise should not only be current but also pioneering, particularly in the realms of artificial intelligence, machine learning, and advanced imaging techniques — all of which are pivotal to revolutionizing steel manufacturing. A partner with a robust research and development wing, staffed with a cadre of adept scientists and engineers, is vital to fostering a culture of continuous innovation and technological breakthroughs.

A history of achievement and collaboration also speaks volumes. Potential partners should bring a dossier of success stories, showcasing their capabilities in steering digital transformation projects to fruition, especially in the realm of manufacturing. The ability to reference past collaborative efforts that have significantly moved the needle in terms of technological or process enhancements would be a strong indicator of their capacity to contribute meaningfully to the project.

Lastly, the infrastructure that these partners have in place cannot be overlooked. For a project as ambitious as SMS Concast's, cutting-edge facilities and equipment are non-negotiable. These resources will be the cradle for the development and testing of the new technologies envisioned. Additionally, an advanced and resilient IT infrastructure is essential to underpin the sophisticated data processing that such digital and AI-driven applications demand.

In sum, the ideal partners for SMS Concast's digital leap are those that not only share a vision for innovation within Switzerland but also bring a proven record of scientific acumen, technical expertise, and the necessary infrastructure to turn visionary ideas into tangible successes.

5. Project Setup

5.1 Project dates

Project duration: 01/12/2023 - 30/11/2024 (12 M)

		start	end	length	Leader	year 1											
						1	2	3	4	5	6	7	8	9	10	11	12
WP1	**Management, quality control, and exploitation**	***1***	***12***	*12*	**CSEM**												
T1.1	Project Management and Admnistration	*1*	*12*	*12*	CSEM												
WP2	**Hardware development, Design and Construction**	***1***	***4***	*4*	**CSEM**												
T2.1	Hardware system design		*2*	*2*	CSEM												
T2.2	Prototype construction	*2*	*3*	*2*	CSEM												
T2.3	Raspberry Pi integration	*3*	*4*	*2*	CSEM												
WP3	**Data Acquisition and Labeling campaigns**	***4***	***8***	*5*	**CSEM**												
T3.1	Design a data collection protocol	*4*	*5*	*2*	CSEM												
T3.2	Execution of the data collection and labeling protocol	*5*	*8*	*4*	SMS Concast												
WP4	**Development of the OCR algorithm**	***8***	***9***	*2*	**CSEM**												
T4.1	Preliminary data and Implementation of the CNN	*8*	*8*	*1*	CSEM												
T4.2	Deployment and Integration to the hardware	*8*	*9*	*2*	CSEM												
WP5	**Tests , Validation and Future Steps**	***9***	***12***	*4*	**CSEM**												
T5.1	Testing in representative conditions	*9*	*11*	*3*	SMS Concast												
T5.2	Performance Evaluation Metrics	*11*	*12*	*2*													
T5.3	Error Analysis	*11*	*12*	*2*													

5.2 Work packages and milestones

Work Package: WP1 Management, and exploitation

Work Package dates

Starting month	1
Duration	12
End	31/11/24
Time allocation per partner in hours	CSEM - Centre suisse d'éléctronique et de microtechnique: **150**

Description

This work package is dedicated to the overall coordination of project activities, quality control, cost/budget management, and exploitation of results.

Activities

	Activity	Responsible	Deliverable/Milestone
T1.1	Project Management and Administration	CSEM	Final Report
	Ensure that the project plan is followed according to the timeline and budget, monitor technical execution and deliverables, coordinate technical development		

Work Package: WP2 Hardware development, Design and Construction of a cost-effective hardware system

- Incorporate a single integrated camera with built-in RGB LED flashes
- Raspberry Pi to optimize energy consumption in the OCR system

Work Package dates

Starting month	1
Duration	3
End	29/02/24
Time allocation per partner in hours	CSEM - Centre suisse d'éléctronique et de microtechnique: **250**

Description

The Hardware Development work package focuses on designing and constructing a cost-effective hardware system with specific features to enhance OCR performance. The key components include a single integrated camera and built-in RGB LED flashes for improved image capture and illumination control. The integration of a Raspberry Pi is employed to optimize energy consumption within the OCR system, leveraging the Pi's energy-efficient characteristics and computational capabilities. This combination enables precise control over lighting conditions through RGB LED flashes, contributing to accurate character recognition. Additionally, the Raspberry Pi ensures energy efficiency, making the overall system both effective and economical. The work package aims to create a streamlined and energy-conscious hardware solution to enhance OCR accuracy and reduce operational costs.

Activities

	Activity	Responsible	Deliverable/Milestone
T2.1	Hardware system design	CSEM	Hardware design document, system specifications and preliminary design prototype
	Develop detailed specifications for the cost-effective hardware system, outlining the integration of a single camera with built-in RGB LED flashes. Create a design plan that ensures optimal functionality.		
T2.2	Prototype construction	CSEM	Functional hardware prototype, initial test results and a report on any issues encountered
	Assemble the prototype of the hardware system, incorporating the selected camera and RGB LED components. Test the individual elements for functionality and compatibility.		
T2.3	Raspberry Pi Integration	CSEM	Functional hardware system
	Integrate the Raspberry Pi into the hardware system to optimize energy consumption.		

Work Package: WP3 Data acquisition and labeling campaigns

Work Package dates

Starting month	4
Duration	5
End	31/07/24
Time allocation per partner in hours	CSEM - Centre suisse d’éléctronique et de microtechnique: **100** SMS Concast: **700**

Description

In the work package, images of machine-generated handwritten digits will be acquired and labeled to train the OCR algorithm. The hardware system assembled in WP2 will be used to collect these images. A GUI for accelerating labeling time will also be developed. The goal is to label 70,000 images

Activities

<table>
<tr><td></td><td>Activity</td><td>Responsible</td><td>Deliverable/Milestone</td></tr>
<tr><td rowspan="2">T3.1</td><td>Design a data collection protocol</td><td>CSEM</td><td>Data collection and labeling protocol</td></tr>
<tr><td colspan="3">This protocol will describe the procedure for image collection and annotation. It will set up the pipeline of for data extraction and a pipeline for the GUI (one could think of an interface in which displays images one by one and that allows the user to position labels where the digits are.</td></tr>
<tr><td rowspan="2">T3.2</td><td>Execution of the data collection and labeling protocol</td><td>CSEM</td><td>Labelled dataset of billet images</td></tr>
<tr><td colspan="3">Billet images will be collected and annotated using the protocol described in T3.1</td></tr>
</table>

Work Package: WP4 Development of the OCR algorithm

Work Package dates

Starting month	9
Duration	2
End	30/09/24
Time allocation per partner in hours	CSEM - Centre suisse d’éléctronique et de microtechnique: **300**

Description

The Development of the OCR algorithm work package focuses on creating a robust digit classification system, emphasizing the use of Convolutional Neural Networks (CNNs) for accurate identification of machine-generated handwritten digits. It includes preprocessing the labeled dataset and training a CNN algorithm following a standard pipeline, incorporating input layers, convolutional layers, pooling layers, fully connected layers, and an output layer. Overall, the work package aims to leverage CNN capabilities to develop an effective digit classification system using a well-structured dataset and algorithmic approach.

Activities

<table>
<tr><th></th><th>Activity</th><th>Responsible</th><th>Deliverable/Milestone</th></tr>
<tr><td rowspan="2">T4.1</td><td>Preliminary data analysis and Implementation of the CNN</td><td>CSEM</td><td>Statistical analysis report of the datasets with a preliminary feasibility study of the ML modeling

CNN Algorithm</td></tr>
<tr><td colspan="3">Implement a CNN algorithm and optimize by iterating to find the appropriate optimizer and loss function</td></tr>
<tr><td rowspan="2">T4.2</td><td>Deployment</td><td>CSEM</td><td>CNN integrated into the hardware system to provide real time classification</td></tr>
<tr><td colspan="3">The algorithmic pipeline described above will be compiled and optimized for real time performance. The AI algorithm and the hardware will be combined</td></tr>
</table>

Work Package: WP5 Tests, Validation and Future Steps

Work Package dates

Starting month	10
Duration	2
End	30/11/24
Time allocation per partner in hours	SMS Concast: **200** CSEM - Centre suisse d’éléctronique et de microtechnique: **100**

Description

This work package focuses on rigorously evaluating the developed OCR algorithm's performance and validating its effectiveness in accurately classifying machine-generated handwritten digits. This phase is crucial for ensuring the algorithm's robustness, generalization, and reliability under diverse conditions. The testing and validation process involves assessing the model's accuracy, precision, recall, and other relevant metrics using a reserved dataset. Additionally, the work package aims to identify potential areas for improvement and fine-tune the algorithm for optimal performance.

Activities

	Activity	Responsible	Deliverable/Milestone
T5.1	Testing in representative conditions	SMS Concast	Dataset
	Prepare a separate testing dataset distinct from the one used during training, collected under real conditions using the new hardware		
T5.2	Performance evaluation metrics	CSEM	Performance metrics report, highlighting strengths and weaknesses Error Analysis report Final report, with recommendations for further algorithm and system improvements
	Evaluate the algorithm's performances under real-world conditions using metrics such as accuracy, recall, and F1 score		
T5.3	Error Analysis	CSEM	Error Analysis report Final report, with recommendations for further algorithm and system improvements
	Conduct a detailed analysis of OCR error, identifying common misclassifications and potential causes		

5.3 Risk Management

See attachment (FMEA table)

5.4 Planned hours

Total planned hours per partner

Responsible Organisation	Planned hours
SMS Concast	1000
CSEM - Centre suisse d'électronique et de microtechnique	800
Total Planned hours	**1800**

Number of the total planned hours dedicated to project management activities:

150 hours (7.5%)

Explanation

The management tasks are already included in WP1 of the work plan (Management and exploitation)

5.5 Project Budget

Financial Overview

	Research Partners	Total Contributions
Salary costs	119’040.00	
Material costs	192.00	
Total contribution needed		**119’232.00**

Salary costs

CSEM - Centre suisse d’électronique et de microtechnique

	Hourly rate (CHF)	Time on project (h)	Amount (CHF)
Project Manager	74.00	125 h	11’100.00
Research partner	74.00	50 h	3’700.00

representative			
Technical Team	60.00	250 h	15’000.00
Risk Management Consultant	60.00	25 h	1’500.00
Data Analysis Experts	70.00	350 h	24’500.00
Salary costs			55’800
Employer’s contributions	20%		11’160.00

Total Salary costs **66’960.00**

SMS Concast

	Hourly rate (CHF)	Time on project (h)	Amount (CHF)
Head of supply chain	74.00	50 h	3’700.00
Technical Team	74.00	50 h	3’700.00
Market and Business Strategy Consultant	40.00	900 h	36’000.00
Salary costs			43’400
Employer’s contributions	20%		8’680.00

Total Salary costs **52’080.00**

Material Costs

Material costs that are necessary for the proper execution of the innovation project can be added below. All amounts can be listed with VAT. The respective costs for VAT must be included in the requested amount. Innosuisse does not pay costs for publication of research results, the use of research infrastructure acquired by third-party funds explicitly provided for this purpose and travels within Switzerland.

Description	Explanation	Category	Used by	Amount
Raspberry Pi Link to product	Series of small, single-board computers developed by the Raspberry Pi Foundation.	Apparatus	CSEM, SMS Concast	87.00
Arducam Pi Camera Link to product	Arducam 16MP Pi Camera 4K, IMX298 for Raspberry Pi Camera, MIPI Camera Module, Plugged into Native MIPI CSI-2 Port on Raspberry Pi.	Apparatus	CSEM, SMS Concast	70.00
RGB LED Link to product	LED Panel Mount Indicators Signal Lamp SMD LED 5mm RGB	Apparatus	CSEM, SMS Concast	35.00

Total Material costs **192.00**

6. Project results on Data
(16’070 characters)

6.1 Methodology / Reasoning

6.1.1 Explain reasoning behind what you did.

The Machine learning part of this project aims to develop a robust Optical Characters Recognition system (OCR) using a digit classification algorithm. The goal is to build a model that takes pictures of semi-finished casting products as input and then can output the text written on them, which is actually their Id. It is crucial to identify it properly in order to get a good logistic tracking system which is very important for traceability of the semi-product sold, its nature and all its relevant properties. We know that the previous version, built by CSEM in 2011, used for the digits classification achieved a 98% recognition accuracy on the digits, which is already pretty high. But, regarding the new mathematical models that have emerged in computer vision, we decided to explore the opportunity to eventually improve this accuracy score, which could result in significant cost savings and product quality improvement for SMS Concast.

For this project, multiple steps are necessary. First, the data collection and appropriate labelling will provide the necessary dataset for the training part. Then, we need to put the data in the appropriate format to be readable by a Machine Learning algorithm. After that, it is very important to build the model architecture, inspired from recent innovations and recent state of the art algorithms for similar tasks, that is key to the success of this project. We can also design some data augmentation processes to improve the model robustness to new unseen data and its understanding of the characters. A very important task is to select the right hyperparameters. If well chosen, they can significantly improve the success rate of our character classification system. Once the model has attained an appropriate training performance, we must test its result on a new unseen test set. It is very important to prevent overfitting. Once the model has achieved required test performance, we can save its weights and use them in the industrial process.

After that, we must think of the MLOps part. It includes the data preprocessing, the loading of the model, the classification of the digits on each picture, and then saving the results into the appropriate format. We must ensure our model classifies our digits in a reasonable time.

6.1.2 Data collection

For the collection of the images, we must take some pictures of the steel blocks at some point in the production line. It is required to use some customized hardware that know when a steel block is at the right position and that it's time to take a picture of it. Each steel block must be associated with an image of it that has been taken. The picture is colored. The dimension of the characters pictures is 64 x 32 x 3 (height, width, number of colors). We obtained a dataset of roughly 75'000 characters images to train our model on.

6.1.3 Labelling

For the labelling part, we must do it manually for a lot of images, in order to get a n important number of training data. This includes the labelling of the characters we have, their order, and finally their position in the image. The aim here is to have each image associated with a list of characters from this set of unique characters {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, C, E, X} with length 11 (each id is composed of 11 characters). Each character has to be associated with a position and to do so, we must define a (x, y) origin, axis, and scale. We decided to put the origin of the (x, y) axis at the top left corner of each picture, and the y axis goes from the top to the bottom along the height of the image, whereas the x axis is along the width of the image from the left to the right. The scale of the (x, y) axis is counted as the number of pixels. Once we have every character and their (x, y) position, we define their order as their position id in the list of the corresponding image. In fact, we want the memory used to be minimal and don't need to provide an additional list of rank to do so.

6.1.4 Which algorithm / pipeline

For the algorithm/pipeline, we decided to divide it into several steps

1. The training part:

We have to include the data loading into the right format of (number of images, height, width, number of colors) matrix and appropriate memory management regarding the RAM capacity. We decided to perform a train/test split of the data using the Sklearn *train_test_split*function. We set a train set size of 60'000 images, which provides a test set of roughly 15'000 images. We also change the data type to float and normalize it by applying a division by the pixel maximum value (255) on each of them. We also have to assign each character of the possible labels with a unique integer. We do that in order to process the label as a vector. We map each integer label of our dataset with the same integer and the remaining characters are mapped according to {C:10, E:11, X:12}. This mapping is necessary because we need to transform the labels into the appropriate target vector using Keras *to_categorical* function.

Then we must create the machine learning model architecture. We decided to use a Convolutional Neural Network architecture that is summarized in the table below. Convolutional Neural Networks are to date the bests models to work with images and to capture spatial information by applying learnt convolutional operators through the image to capture specific shapes, regardless of where they are positioned. Before to obtain this specific architecture below, we started from another one, which is very similar and that is used to classify one dimensional MNIST handwritten digits. We decided to try different architectures hyperparameters and the results of these tests are described on the results part. Here, the kernel applied in the convolutional layers are of size (3, 3).

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 62, 30, 64)	1792
batch_normalization_5 (BatchNormalization)	(None, 62, 30, 64)	256
activation_6 (Activation)	(None, 62, 30, 64)	0
conv2d_5 (Conv2D)	(None, 60, 28, 64)	36928
batch_normalization_6 (BatchNormalization)	(None, 60, 28, 64)	256
activation_7 (Activation)	(None, 60, 28, 64)	0
max_pooling2d_2 (MaxPooling2D)	(None, 30, 14, 64)	0
conv2d_6 (Conv2D)	(None, 28, 12, 128)	73856
batch_normalization_7 (BatchNormalization)	(None, 28, 12, 128)	512
activation_8 (Activation)	(None, 28, 12, 128)	0
conv2d_7 (Conv2D)	(None, 26, 10, 128)	147584

batch_normalization_8 (BatchNormalization)	(None, 26, 10, 128)	512
activation_9 (Activation)	(None, 26, 10, 128)	0
max_pooling2d_3 (MaxPooling2D)	(None, 13, 5, 128)	0
flatten_1 (Flatten)	(None, 8320)	0
dense_2 (Dense)	(None, 512)	4260352
batch_normalization_9 (BatchNormalization)	(None, 512)	2048
activation_10 (Activation)	(None, 512)	0
dropout_1 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 13)	6669
activation_11 (Activation)	(None, 13)	0

==

Total params: 4530765 (17.28 MB)

Trainable params: 4528973 (17.28 MB)

Non-trainable params: 1792 (7.00 KB)

The loss we use for the learning of this model is the *categorical_crossentropy* loss from the Keras library. It can be seen as minus the log of the probability of the actual label. The cross-entropy loss is defined as follow:

$$CE = -\sum_{labels\ i} t_i * log\big(f(s_i)\big)$$

With:

- t_i = 1 if i is the actual label else 0
- f(s_i) = p(t_i=1| image data)

The optimizer we use for the training is *Adam*, which is set with a learning rate of 1e-3, beta1=0.9 and beta2=0.999, epsilon=1e-07.

We also define a suitable data augmentation process to make the trained model more robust and generalize better to new unseen data. This algorithm allows to slightly modify the images randomly by applying multiple operations on them to make the Machine Learning model learn on more “artificial” data. The different operators are the rotation of the images, shifts along width or height, shear, zoom, changes in brightness, and other operators. To do perform this data augmentation, we used the Tensorflow Keras built-in function to do automatic data augmentation: *ImageDataGenerator*. The advantage of this function is that it takes the training images and perform the images modification “on-the-fly”, meaning that the transformation is done as the training goes on and not previously to it. It allows for a much better management of the memory and avoid creating a huge dataset from the initial one. Here is how we set the parameters for our data augmentation process:

- rotation_range=8 — Int. Degree range for random rotations.
- width_shift_range=0.08 — Float: fraction of total width
- shear_range=0.3 — Float. Shear Intensity
- height_shift_range=0.08 — Float: fraction of total height
- zoom_range=0.08 — Float or [lower, upper]. Range for random zoom.

Here is a list of the remaining hyperparameters we choose:

- batch size = 128
- number of epochs = 15

Once the model has been trained, tested, and has proven to have the required results, we can save its weights to use it later.

2. The Operational part:

Once the model has been learned and saved with appropriate results, we can use it whenever needed. For this part, which is more about “MLOps”, we use the same data loading and preprocessing steps as for the training part, unless we don’t separate into train/test sets of course. Then we load the model from its saved weights, and it is ready for use. Then for each picture taken in the industry line, we process the data with the Machine Learning algorithm and output the digit that has the highest probability. After that, we map our integer to its unique character. We do this process for every character of each image and get the id of each cast steel.

6.2 Results

6.2.1 What did not work, why?

For the hyperparameters selection, we decided to start from an architecture that is known to work well for the MNIST dataset. This is the CNN architecture used by "Introduction to Deep Learning with Keras and TensorFlow" of the data science for business course at EPFL.

We first tried, with all the other hyperparameters constant, to change the kernel size. We tried (4,4) and (5,5) but we observed that the test accuracy decreased as the kernel size was increasing. We then decided to keep it at (3,3).

We then tried to vary the number of different kernels at each convolutional layer of the model. By augmenting this number by a multiple of 4 (layer with 32 kernels is changed to have 4x32=128 different kernels) for each layer we saw a decreasing in the model test accuracy. But, using a multiple of 2 we saw an improvement. We decided to keep it with a multiple of 2. This means that each convolutional layer initially with a number of different kernels set to 32 changed to 64 and those with 64 to 128.

6.2.2 What worked

We also decided to try different numbers of epochs to train the model on. Initially we did a small number of epochs (5) and saw that number of epochs that gives the best test accuracy was 3. But, when we restarted from scratch the learning procedure, we saw that the resulting accuracy was not optimal and that it was at the 4^{th} epoch. We decided to try then with 15 epochs to see what happened. We saw that the testing accuracy was not always increasing epochs after epochs but that on average It was. We understood that the learning process is noisy and that we don't get improvement at each epoch. After 15 epochs we get the best results and decided to keep it like that as the improvements were almost invisible after all these learning steps.

After all those trials described above, we finally get our best model and obtained the following results:

- Train accuracy = 99.987 %
- **Test accuracy = 99.933 %**
- Train loss = 5.9973e-04
- Test loss = 0.0033

6.2.3 Evaluation and reasoning for this evaluation

For the evaluation of our model, we refer to the test accuracy as the performance of our model. It is the percentage of digits that are classified in the right section. Our test accuracy of 99.933 % means that on 10'000 images, we make on average 6,7 mistakes, whereas the previous model (98% accuracy) made 200 mistakes. Hence, the number of mistakes has been divided by almost 30.

We must be careful with this test accuracy as the model has been selected on this metric. It is possible that the model then gets slightly worse results when used on new data. We call that the generalization error. In fact, we optimized our model by trying manually to get the best test results, which create a bias toward the real generalization error.

6.4 Future work

Here are some further steps to eventually improve this model performances:

- Try to optimize the Adam optimizer parameters (lr, Beta1, Beta2). This is done by testing manually different values for the parameters.

- Try different optimizers. Adam is one of the most popular optimizers for Machine Learning training, but there are plenty of them, like SGD with momentum or RMSprop. One can try them and assess if they improve the model performance.

- Try learning rate scheduling, which is a very popular method to avoid overfitting. It consists of progressively decreasing the lr during the training procedure.

- Vary the dropout rate. It is set to 0.2 in our case, but maybe some other value would lead to better results.

- Try gradient clipping. It is a method that prevents too high gradient from backpropagation by applying a clip operator on the gradient before processing it by the optimizer.

- Check if the data labels are imbalanced and assess this problem. This can cause some bias in the learning procedure with a model that will predict more often the dominant labels as a result. We can assess this problem by rebalancing the dataset through data augmentation, or by creating a bias in the loss function with weights that gives more importance to the rarest labels.

- Try to make the Convolution Neural network deeper with additional convolutional layers.

- Try the new transformer-based architecture for computer vision tasks: Vision Transformer (ViT). It is a very good model that achieves 98%+ accuracy on the MNIST dataset. It could be very good in our use case.