Integrating Activities for Advanced Communities

D6.4- Report on Future Strategy and Planning for the Area of AI and ML to be Applied in Arctic Research

Project No.871120– INTERACT

H2020-INFRAIA-2019-1

Start date of project: 2020/01/01 | Duration: 48 months

Due date of deliverable: 2022/08/31 | Actual Submission date: 2022/08/20

Lead partner for deliverable: AFRY
Author: Maria Erman, Markus Skogsmo, Carl Sundström

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the Consortium (including the Commission Services)	
CO	Confidential, only for members of the Consortium (including the Commission Services)	

Table of Contents

Publishable Executive Summary ... 4
Table of Acronyms ... 5
1. Introduction ... 6
1.1. Background ... 6
1.2. Purpose and Goals ... 7
1.3. Outline ... 7
2. Theoretical Framework ... 8
2.1. An Introduction to Artificial Intelligence and Machine Learning ... 8
2.2. An Introduction to Artificial Neural Networks ... 11
3. Conducted Work for WP6 ... 13
3.1. Deliverable 6.1 - Pre-study on Inquiries and Needs from Station Managers and Researchers ... 13
3.2. Deliverable 6.2 - Workshop with Demonstration on AI and ML Technologies ... 13
3.3. Deliverable 6.3 – Demonstration on Using Machine Learning on Example Data ... 14
3.3.1. Computer Vision for Classifying Animals and Mushrooms ... 14
3.3.2. Satellite Data for Analyzing Landscape Changes ... 14
3.3.3. Optical Character Recognition and Natural Language Processing ... 15
3.3.4. Cloud Computation in Connection with Machine Learning ... 15
3.3.5. Summary of the Demonstration ... 15
3.4. Image Augmentation to Create Lower Quality Images for Training a YOLOv4 Object Detection Model ... 15
3.4.1. Introduction ... 15
3.4.2. Tasks and Scope ... 15
3.4.3. Fundamentals of Digital Image Processing ... 16
3.4.4. Results and Conclusions ... 19
3.5. Automated Digitization and Summarization of Analog Archives: Comparing Summaries made by GPT-3 and a Human ... 19
3.5.1. Introduction ... 19
3.5.2. Tasks and Scope ... 20
3.5.3. Purpose and Goals ... 21
3.5.4. Results and Conclusions ... 21
3.6. Computer Vision for Camera Trap Footage: Comparing classification with object detection ... 23
3.6.1. Introduction ... 23
3.6.2. Tasks and Scope ... 24
3.6.3. Results and Conclusions ... 25
3.7. Searching and Recommending Texts Related to Climate Change ... 26
3.7.1. Introduction ... 26
3.7.2. Tasks and Scope ... 27
3.7.3. Results and Conclusions ... 27
3.8. Deep Learning for Iceberg Detection in Satellite Images ... 28
3.8.1. Introduction ... 28
3.8.2. Tasks and Scope ... 29

3.8.3. Results and Conclusions 29
3.9. Collaboration with INTERACT III Research Stations 29
4. Future Strategy and Planning 31
4.1. Opportunity Analysis 31
4.1.1. Potential Use Cases of AI Technology 31
4.1.2. Open-Source Solutions 31
4.1.3. Commercial Solutions 32
4.1.4. Potential Methodologies for AI Applications 32
4.2. Future Needs of AI Technology 34
4.3. Future Goals 34
4.4. Identified Initial Steps 35
4.5. Identified Challenges 35
4.6. Suggested Roadmap 36
5. Summary and Conclusions 37
5.1. Summary 37
5.2. Conclusions 37
Bibliography 41

Publishable Executive Summary

This document presents the insights gained from previously undertaken work by work package 6 (WP6) in INTERACT III. In addition, it outlines future strategies and planning for artificial intelligence (AI) and machine learning (ML) applied in Arctic science.

As such it presents condensed and concise contents from previous WP6 deliverables, which have consisted of workshops, a pre-study on requirements and needs from station managers and researchers in the fields of AI and ML, as well as the outcome of a demonstration of previously conducted pilot projects. This document attempts to put all the work in context, in order to examine how its various parts are connected, as well as what conclusions can be drawn from the work as a whole.

Moreover, in-depth and technically advanced master thesis work has been conducted by skilled students in the areas of computer vision (CV) and natural language processing (NLP) under INTERACT III, supervised by AFRY and in connection with Uppsala University. Summaries and conclusions are all collated in this document in the topics of performance of object detection models handling simulated, lower quality photographs; using CV models on camera trap footage; deep learning (DL) for iceberg detection in satellite images; ML models summarising digitised old logbooks; as well as an examination of recommendation system models trained on 10000 abstracts on climate change.

Finally, future strategies on enabling opportunities to adopt a data driven approach for Arctic research is presented. As a first step, it is imperative to understand what the data driven maturity is in the organisation. Furthermore, it is proposed to examine and make a list of manual and resource consuming tasks in the organisation, as well as framing or translating them into a problem interpretable by an AI model to ensure proper implementation. For this, added competence and increased collaboration with actors with a data driven perspective is recommended, e.g., even a PoC undertaken by a master thesis student can go a long way.

Table of Acronyms

AGI	Artificial general intelligence
AI	Artificial intelligence
ANI	Artificial narrow intelligence
ANN	Artificial neural networks
API	Application programming interface
CV	Computer vision
DL	Deep learning
ML	Machine learning
NLP	Natural language processing
OCR	Optical character recognition
SC	Societal challenges
WP	Work package
YOLO	You only look once

1. Introduction

1.1.Background

With rising temperatures occurring at least twice as fast as in the rest of the world, the Arctic is an important area to study in order to gain an in depth understanding of climate changes. It is not only a region, but also a physical, biological, chemical and climatological system. No other area is as impacted by global warming as the Arctic. The region may seem remote, but with ice melting, permafrost thawing, retreating glaciers, and consequent rising sea levels, it also affects the climate and the weather elsewhere on earth [1].

Global warming directly and indirectly impacts plants and animals, with for instance species such as polar bears, walruses and Arctic foxes finding it more difficult to hunt with diminishing ice. As an example of cascading effects, it can be mentioned that lichens and mosses, which are food sources to Caribous, are sensitive to warming, hence leading to a declining Caribou population. This in turn also affects scavengers and predators [1][2].

Indigenous people, who have lived in the Arctic for thousands of years, and whose cultures are shaped by the environment, are particularly percipient observers regarding climate changes. Noticeably, indigenous people have observed a reduction of sea ice, weather patterns changing and becoming less predictable, and species not previously seen appearing etcetera [3, pp. 4–5][4].

Environmental change in the Arctic is thus highly complex, and combined with few people living in the Arctic, data is sparse. Hence, to aggregate data, and attempting to model and understand the effects of environmental change on ecosystems, it is of paramount importance to conduct field work in a variety of areas, with the aim of obtaining an understanding of what the parameters are and how they are related. Furthermore, various forms of documentation, be it measurements, pictures, videos, sounds, or satellite imagery, will be required, not only for present, but also future research.

Great progress has been made by Academia in analyzing and documenting environmental change, and with the recent explosion of artificial intelligence (AI) and machine learning (ML) techniques, there are now new possibilities and perspectives to consider. It has become accepted that “there are many ways of knowing”, and in addition to conventional and traditional knowledge, there are also resources such as logbooks, expedition reports, photographs and paintings that can be used to extend existing knowledge.

In essence, AI is about mimicking human intelligence, with ML being a set of tools, techniques, and algorithms to achieve AI. Moreover, AI and ML have been recognized not only as powerful research tools in several scientific disciplines, but it also shines when it comes to automating manual and tedious work already performed by humans today. It does this by learning from sources of information and applying this knowledge on new information. As such, application areas are as diverse as image recognition, object detection, self-driving cars, optical character recognition (OCR), syntax analysis, sentiment analysis of text etcetera [5].

1.2.Purpose and Goals

Valuable data are collected by dedicated researchers involved in INTERACT III, and some research stations have conducted measurements for many decades. Not all data collected thus far may seem to have a purpose. However, in the future, analysis of these data may provide the missing link to understanding key concepts and bringing crucial insights into climate change research. Those additional photographs, recordings or measurements could aid e.g., a future state-of-the-art AI algorithm in finding patterns or discover previously unknown relationships between parameters. With this in mind, it is important to take into account that even today there are ways for researchers to prepare and curate their data sets to be as useful and future proof as possible.

The purpose of this report is twofold, i.e., to present summaries of conducted work having been undertaken for WP6, as well as to describe future strategies and how to plan for AI and ML to be applied in Arctic research.

1.3.Outline

Section 1 presents a background, as well as the purpose and goals of this report. Section 2 provides a theoretical framework, in which AI and ML concepts are introduced, such as computer vision (CV), natural language processing (NLP) etcetera. Further, Section 3 is a compilation of previously conducted work of work package 6 (WP6) that encompasses deliverables and master theses conducted under INTERACT III. Section 4 delves into future strategy and planning. Thereafter, Section 5 provides a summary and conclusions for this deliverable and for WP6 as a whole.

2. Theoretical Framework

2.1. An Introduction to Artificial Intelligence and Machine Learning

The foundation of AI as it is currently known, was already laid in the 1950, even with advanced techniques such as artificial neural networks (ANN) being put forth. The hype and interest in AI, however, did not explode until recently. The reason for this is the large amount of data needed to be processed. For example, for a neural network to be able to classify images, thousands of images might be needed to train the network. In the case of text, the equivalent of a thousand books might be needed for a satisfactory result. In other words, a limited amount of computational power and memory needed for processing huge amounts of data prevented the breakthrough of AI until recently.

Two important questions are: "how can we start benefitting from these powerful techniques today?", and "what are the pre-requisites for using these techniques, and in what applications do they shine?". Before answering these questions, a good place to start is to clarify what AI is, what it can and cannot do, and to cut through the hype.

AI has several connotations, and one of many ways to approach understanding AI, is to divide it into artificial narrow intelligence (ANI) and artificial general intelligence (AGI), see Figure 1. There is a great divide between these concepts. While ANI does one thing incredibly well such as object detection, image classification, text analysis, translation etcetera; in contrast, AGI is about emulating humans as a whole, including human intelligence, and even surpassing humans. Most notably, there has been very little progress with regards to AGI, while ANI is thriving and creating immense value [6].

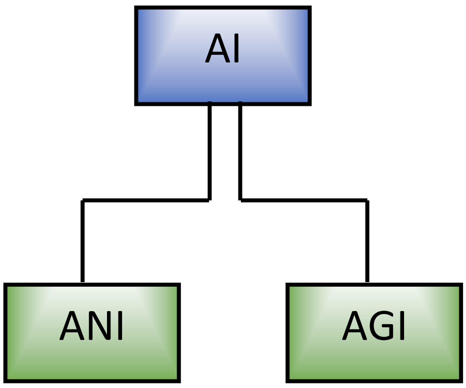


Figure 1: AI divided into two categories, artificial narrow intelligence and artificial general intelligence.

AI encompasses both ANI and AGI, with the former concept being what is used professionally with tangible results. On the contrary, in pop culture and literature, it is AGI that inspires popular myths and stories - after all, AGI does render itself very well into a dramatic and immersive format. The progress of ANI is thus mixed up with AGI, leading to irrational fears concerning the rise of the machines, robots taking over both existentially and/or rendering humans useless.

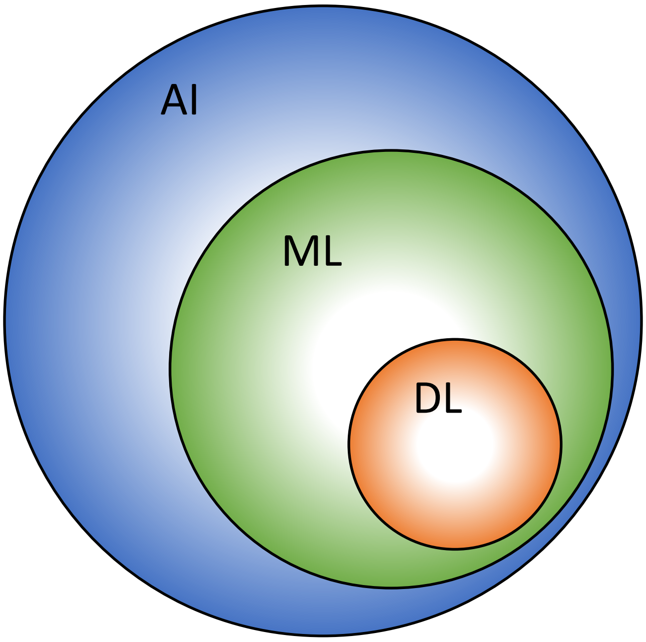


Figure 2: Artificial intelligence, of which machine learning and deep learning are conceptual subsets.

Furthermore, terms such as ML and deep learning (DL) are often used in connection with AI. There is confusion with regards to how to differentiate these concepts, and while they are used interchangeably, they do not mean the same thing. The obvious questions are, “what do these terms mean”, and “how do they relate to each other?” In short, AI is a superset of ML, which in turn is a superset of DL, see Figure 2.

More specifically, AI is the concept of emulating and mimicking human intelligence, behavioural patterns, interactions, sensations etcetera, and is in essence the theory of mind. AI is probably the hardest of the three terms to define since it has taken on different meanings depending on the domain in which it acts.

ML is one approach of reaching AI, by using methods to automatically learn from data without being explicitly programmed to do so. Moreover, ML can predict outcomes from similar data, and subsequently, improving from experience. To gain a clearer understanding, ML algorithms are commonly classified into three categories, supervised learning, unsupervised learning and reinforcement learning [7]:

- In the case of supervised learning, algorithms learn by example. With regards to training data, inputs are mapped to the correct output. Subsequently, the network adjusts itself based on training. After proper training, the network is employed, whereby new, unknown data are being mapped to the correct output in accordance with previous training.
- With unsupervised learning, the algorithm finds structures on its own, without having an algorithm being trained on pre-specified labels beforehand. Effectively, the aim of the unsupervised learning algorithm is to find hidden information and patterns in the input data, with the algorithm grouping unsorted data according to similarities. The algorithm thus categorizes data based on their similarities, differences, and patterns.
- Reinforcement learning can be likened to teaching a dog. A dog explores, and with curiosity takes actions. These actions lead to consequences. If the dog eats healthy dog food, it is a good dog. On the other hand, if the dog eats the family lasagne, it is a bad dog. The dog learns by rewards and corrections, which is essentially a metaphor for how reinforcement learning works.

DL is the notion of employing ML techniques by using ANNs [8]. It can be said that DL are ML techniques that teach machines to learn by example. In DL, a computer model learns to make classification tasks from a large amount of data in the form of images/videos, text, numbers or sounds [9].

2.2.An Introduction to Artificial Neural Networks

DL is generally concerned with ANNs, which are a technology inspired by how a brain operates with its biological neural network, and as such ANNs aim to mimic how a brain finds patterns or relationships among vast amounts of data [10].

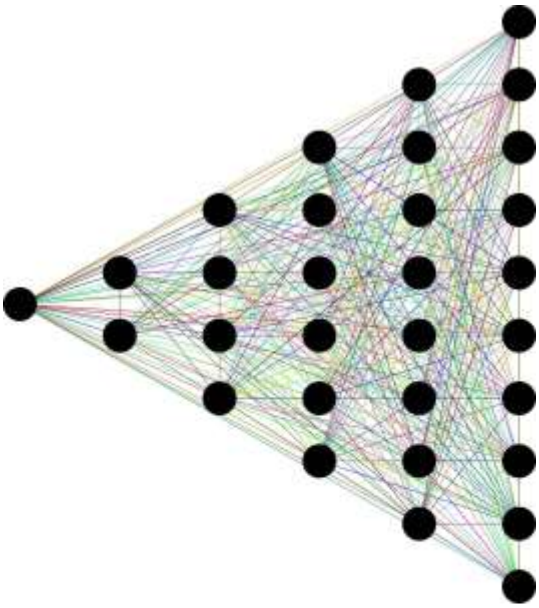


Figure 3: Artificial neural network with interconnected neurons.

An ANN consists of interconnected neurons or nodes structured in layers: An input layer, an output layer and what is denoted as hidden layers that exist between the input and the output layers, see Figure 3. Each layer extracts a different set of features from the input data, where an input, for instance, might be an image, and a set of features being edges. With every layer, the feature detection becomes more and more refined [11].

For a basic overview of how ANNs work, the example of inputs being images and outputs being classes, for instance bird and polar bear, can be used. Firstly, the network is trained, or e.g., a few thousand images are fed into the network. The network then adjusts itself, or the weights of the nodes to be more specific, in accordance with optimization techniques, i.e., we know what images we feed into the network, and we know the output classes. The network, after being properly trained, is now ready for unknown images to be fed into it and then output the correct class, see Figure 4 [12].

Figure 4: An image recognition neural network model detecting animals and classifying them into correct species. (a) Bird. (b) Polar bear.

In a similar way, a neural network can be trained on text. For the network to be accurate, usually text of the order of a few thousand books might be needed. Words, according to the algorithm used, are often mapped to vectors, and after training, patterns can be discerned with similar words being closer to each other, see Figure 5. The results will be highly dependent on the data it was trained on, and in this case the data was trained on a BBC News data set[1] [13].

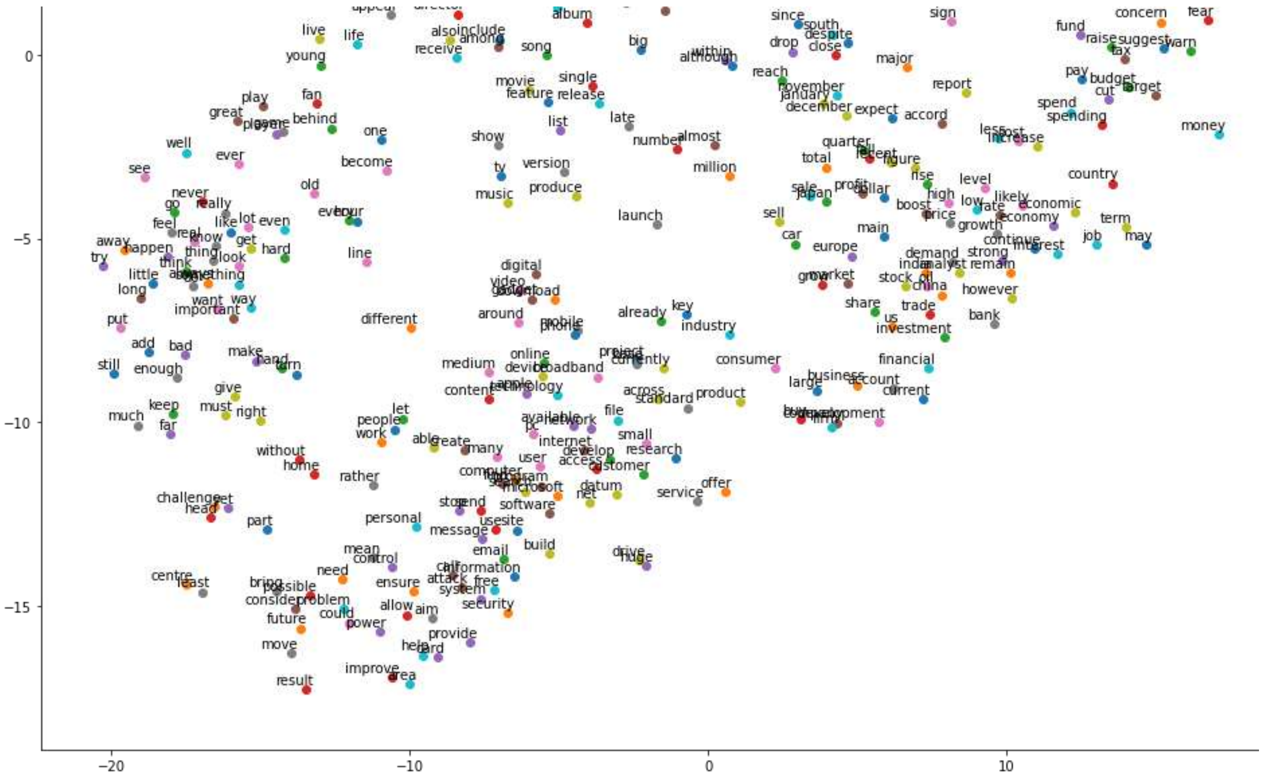


Figure 5: The result of a neural network being trained on text, using a BBC News data set.

Training on huge amounts of data can be problematic, since it is difficult to oversee whether the original data is inherently biased, or worse, there being scathing controversies sparked due to outputs, no matter how rare, being racist, sexist, ageist, etcetera, as was the case with the South Korean chatbot, Lee Luda, trained on 10 billion real-life conversations, and consequently suspended from Facebook for hate speech [14].

ANNs are used in a wide variety of applications, for instance, in time series forecasts, image recognition, classification, self-driving cars, generation of incredibly realistic CGI faces, recommender systems, automatic translation, text to speech, sentiment analysis, etcetera.

[1] The Natural language processing algorithm Word2Vec [13] was implemented, using the Gensim Python library [48].

3. Conducted Work for WP6

AFRY have previously submitted three deliverables, with this report for D6.4, being the fourth and final deliverable. In addition, the work for WP6 has also encompassed master thesis work in the areas of CV and NLP. Hence, sections 3.4-3.8 present summaries of the aforementioned master theses pertaining to AI and ML applied to Arctic research conducted under INTERACT III, supervised by AFRY, and in collaboration with Uppsala University. As such it goes into considerably more technical depth than the reports for the deliverables.

3.1.Deliverable 6.1 - Pre-study on Inquiries and Needs from Station Managers and Researchers

The pre-study, D6.1 [15], presents the results of a combined quantitative and qualitative study on needs, perceptions and uses of AI and ML techniques and methods by station managers and researchers involved in INTERACT III.

The qualitative part of the study consists of a questionnaire provided to the station managers and researchers, as well as structured interviews. This is collated with an analysis of discussions held during workshops to form an idea of how best to proceed with applying AI and ML techniques in INTERACT III. The quantitative part consists of classifying answers into distinct categories to provide a sense of the most important areas of interest for the researchers.

The results of the study showed that several research groups are keenly interested in applying AI and ML in their research but do not know how to do so. The perception appears to be that the main use of AI and ML is to support in decreasing the amount of tedious and error-prone field work and associated analysis and documentation.

3.2.Deliverable 6.2 - Workshop with Demonstration on AI and ML Technologies

This report for the workshop, D6.2 [16], compiles the topics presented during the mini workshop on June 10th, 2020, the workshop held on September 23rd, 2020, as well as the ideas expounded on during the ensuing discussions. Station managers and researchers engaged in the INTERACT III were invited.

The purpose of the workshops was to gain insight into how the participants involved in WP6 could help the INTERACT network, and to discuss ideas for further projects and/or collaborations. Moreover, key ambitions were also to increase awareness of AI and ML, what can be achieved, and pre-requisites for using them. The workshops were conducted virtually via Zoom due to travel restrictions pertaining to the COVID-19 pandemic.

During the workshop, participants gained knowledge of the basics of AI and ML - including its history, how it is presented in popular culture, as well as major areas of AI and ML, such as CV, NLP, and reinforced learning. There were also demonstrations on professional tools in areas such as face recognition, DL and NLP techniques to gain information from and analyzing text.

With regards to the subsequent discussions, participants showed enthusiasm and interest in how AI and ML could be applied in their daily field work to reduce manual work, hence saving valuable working hours that are more efficiently allocated to advanced research. Several researchers present at the workshops expressed that they were new to this field and stated how AI and ML algorithms might benefit their work in Arctic research.

3.3. Deliverable 6.3 – Demonstration on Using Machine Learning on Example Data

The report for D6.3 [17], compiles the topics presented during the demonstration that took place on February 7th, 2022, by WP6. Station managers and researchers of INTERACT III were invited.

One of the goals of WP6 was to raise awareness of AI and ML for researchers and station managers. This includes demonstrating applications and areas, where these techniques excel, their constraints, as well as introducing pre-requisites needed to use them. In addition, a goal of WP6 was indirectly to incite change by showcasing how employing AI and ML techniques does not have to be a huge undertaking, and that small efforts may greatly benefit researchers.

An aim of D6.3 was to demonstrate how ML techniques can be used on example data to showcase specific algorithms and methods, and the outcomes thereof. A wide variety of ML techniques have been used for the demonstration, such as CV, NLP and decision trees in application areas such as classifying animals and mushrooms, using satellite data for earth observations, digitizing and summarizing logbooks, as well as predicting and classifying tabular data.

3.3.1. Computer Vision for Classifying Animals and Mushrooms

Using camera trap surveys combined with CV models, can give ecologists a great tool to improve wildlife monitoring of their stations. By no means perfect, it can still yield consistent and cost-effective results without a complicated setup. The AFRY team recommend interested stations to try it out by purchasing a single camera trap and then processing the images with the MegaDetector [18] model.

There are several ways to utilize CV models and to contribute to their improvement. One way is to join an international crowdsourcing community like iNaturalist [19]. This can be done without any previous experience and still lets you have an intuitive user experience and give powerful results. To get more specific models created for your need, it is often better to train the model yourself. This can result in very precise results for your needs. However, it also requires lots of data. iNaturalist was demonstrated, and the result of the self-trained model was presented on INTERACT research station data.

3.3.2. Satellite Data for Analyzing Landscape Changes

The demonstration also covered how public satellite data can be used to analyze changes in the landscape around the stations. In addition, it was illustrated how pre-existing scripts and ML-results can be used to meet specific needs, e.g., movement and texture changes in glaciers. For more advanced users, scripts could be updated or created from scratch to be especially suited to an uncommon task. However, researchers should start by looking at the pre-existing solutions to quickly get an overview of the possible insights and results suited to their research needs.

3.3.3.Optical Character Recognition and Natural Language Processing

Furthermore, it was described how OCR and NLP can be powerful tools for digitizing old data. It also showed how the free cloud computing service Google Colab [20] can be used as a powerful tool for exploring ML methods.

3.3.4.Cloud Computation in Connection with Machine Learning

The final part of the demonstration showcased how a powerful cloud computation model in Microsoft Azure [21] could be built from scratch without considerable previous knowledge. Azure also provides a graphical UI that visualizes the ML-model, which makes it easier to understand and update the model. Models like these can be trained to make predictions and analyses of different data types, and requires little more than an Internet connection to start developing.

3.3.5.Summary of the Demonstration

To summarize, the demonstration showed that there already exist many exciting possibilities for INTERACT researchers to utilize AI tools to analyze their data and create new insights, even without programming. These possibilities should not be overlooked and the AFRY team highly recommend starting smaller research projects to utilize these techniques.

3.4.Image Augmentation to Create Lower Quality Images for Training a YOLOv4 Object Detection Model

This section is a summary of the work conducted for the master thesis “Image Augmentation to Create Lower Quality Images for Training a YOLOv4 Object Detection Model” supervised by AFRY and in collaboration with Uppsala University. The full master thesis is published in DiVA [12].

3.4.1.Introduction

Data gathered in the Arctic regions are very important for the areas of climate change, geology, meteorology, ecology, glaciology and more. Being able to use already existing data together with more modern, high-quality data may bring even higher performance to predictions in the desired subject.

This project is done to show, in general, the usefulness of AI in Arctic research. Specifically, it is a pilot study to see how older data can be used in modern, state-of-the-art AI algorithms and to:

- Artificially create a data set with introduced disturbances to emulate older images.
- Train a modern object detection model using the created data set.
- Evaluate if the trained model is performant despite being trained on a set of images with disturbances and investigate which disturbances affect it the most.

3.4.2.Tasks and Scope

Data were collected by connecting to a modern data set application programming interface (API) called the Open Image data set v6 [22], containing thousands of images of labelled classes. These images were then augmented to expand the collected data set. Specifically, the augmentations aimed to create images of lower quality, to simulate some properties that might occur in older images, such as lack of color, more noise, and lower quality of contrast and luminescence. This procedure consisted of the following tasks:

- Gathering the data set
- Processing and augmenting the data
- Training a You Only Look Once Version 4 (YOLOv4) object recognition model [23]
- Investigating the performance of the trained model

3.4.3.Fundamentals of Digital Image Processing

Deep neural network algorithms, like most algorithms in AI and ML, need to be trained with a large amount of data to achieve good performance. Image augmentation can be used to build a powerful image classifier using smaller training data and boost the performance of the neural networks. Image augmentation can artificially create training images using different imaging processing techniques such as random translation, rotation, flips, shear, and other geometric transformations, as well as altering the properties such as color, saturation, and luminescence.

It is important to understand how an image is constructed, to be able to perform mathematical operations on it, since the only thing a computer sees are pixel values. There exist different colour spaces (or models), and among them, RGB is the most common way to represent colours digitally and it is a well adopted standard. Each channel in RGB is represented by an unsigned 8-bit number, yielding a decimal value range between 0 to 255 per channel as defined in [24]. An example with the RGB colour channels plotted separately is shown in **Figure 6**.

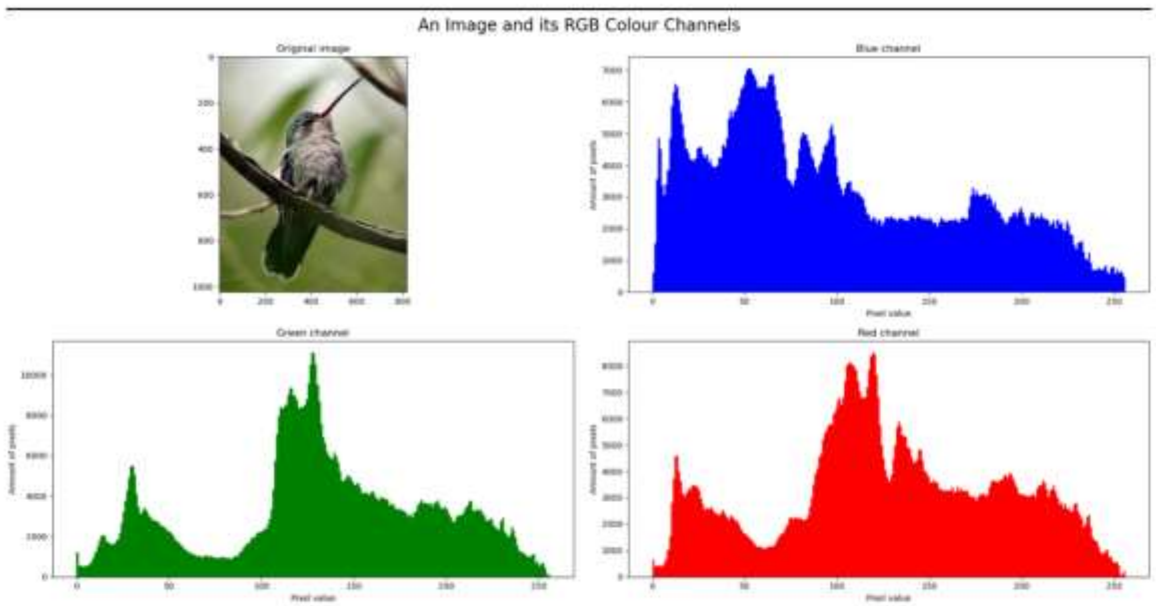


Figure 6: An image's colour channels plotted separately

Older photos are black-and-white images with a grayscale. A grayscale image is one in which each pixel is a single sample that has a value of light intensity. Naturally, these images would be faster to process in an object detector, as the number of values describing the image is a third of the original RGB image and contained in a single colour channel. The histogram of the grayscale version of **Figure 6** is shown in **Figure 7**, where a value of 0 means a completely black pixel, 255 means a completely white pixel, and everything in between is a darker or brighter shade of grey.

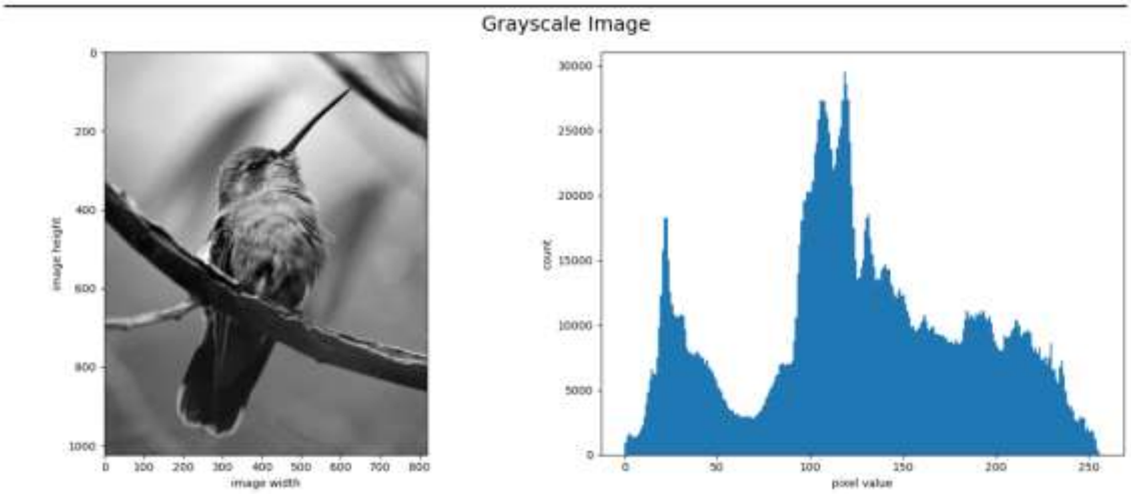


Figure 7: A bird picture converted to grayscale

Noise can appear in an image in many stages, such as acquisition, file format conversion or transmission [25]. Such noise is often Gaussian noise, which is a very common type of noise, especially in relation to images, see **Figure 8**.

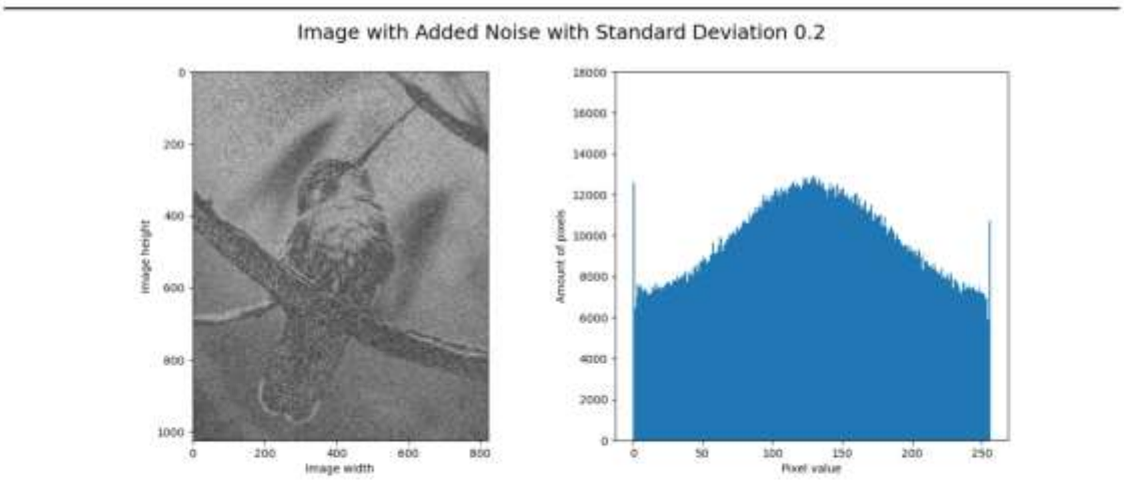


Figure 8: Grayscale image with added Gaussian noise

Gamma correction, also sometimes referred to as Stevens' Power Law [26], is something that an image undergoes several times from when it is captured and finally displayed on a digital monitor. Gamma correcting both image and video is necessary, as the human visual system is well documented to not perceive light in the same linear fashion as it is physically emitted. In this report, it was used to investigate the effect of luminescence and contrast in object detection.

Figure 9 to Figure 11 show the different variations of gamma adjusted images. The pixels in the images with an applied gamma of less than 1.0 are of darker colours, and pixels in the images with applied gamma of greater than 1.0 are of brighter colours, which is displayed by each corresponding histogram.

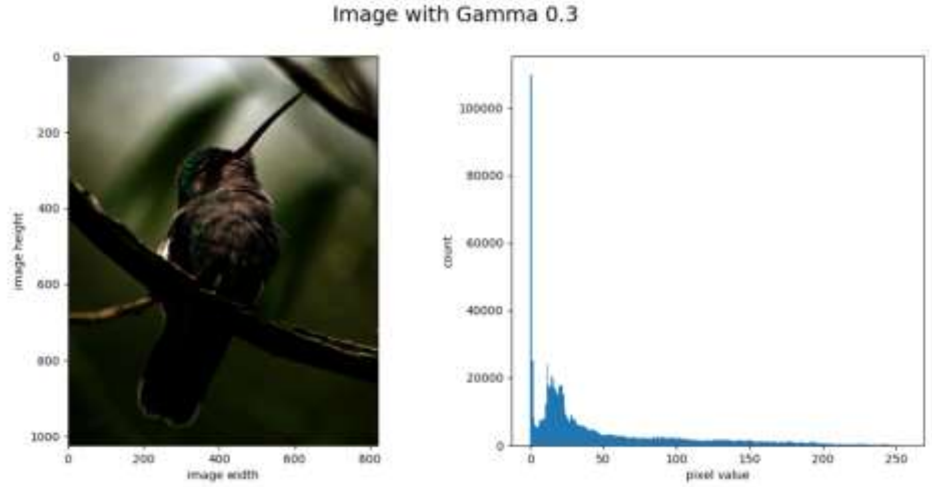


Figure 9: Image with gamma adjustment 0.3

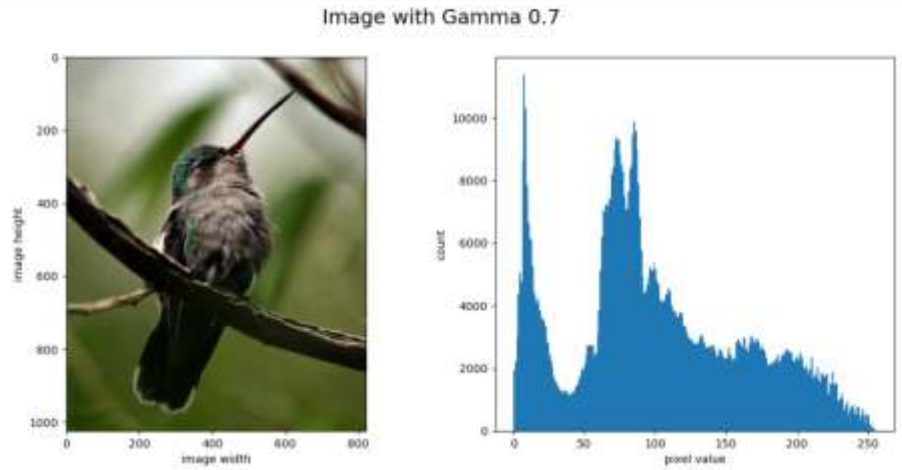


Figure 10: Image with gamma adjustment 0.7

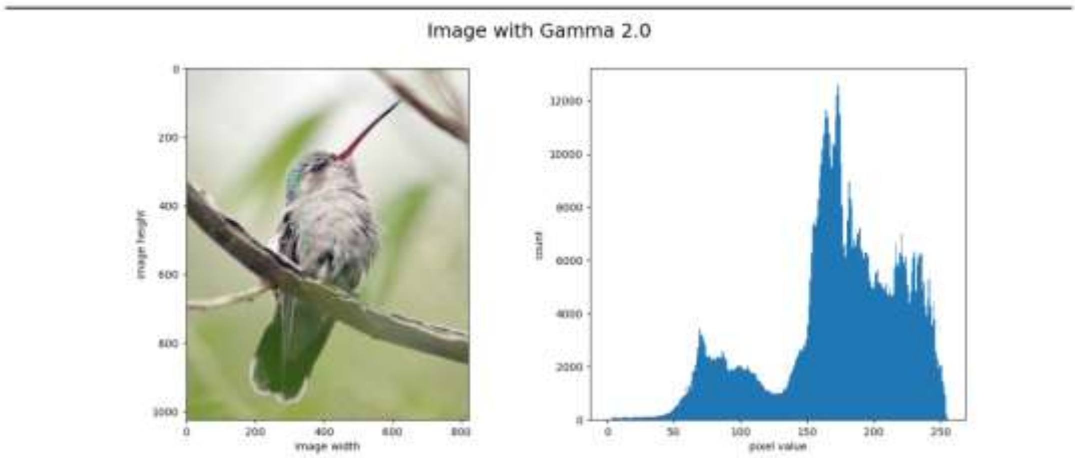


Figure 11: Image with gamma adjustment 2.0

3.4.4.Results and Conclusions

A YOLOv4 object detection model was trained on two classes of images to investigate the performance impacts of disturbances in the training data set. The results showed that even with a fairly small data set, a decent object detector could be produced relatively quickly. Modern object detectors such as YOLOv4 are sophisticated enough to make this available.

An expanded data set was created by augmenting the initial data to contain various disturbances. A model was successfully trained on the augmented data set. Noisy images seemed to be more impacted, i.e., the presence of noise plays a more significant role than some difference in exposure or saturation. The results showed an increased spread and worse prediction results for increasing levels of Gaussian noise, whereas the gamma adjusted images showed very similar results across all variations of gamma values. This could be because of reduced edge detection capabilities, as the edges of objects are blurred out against the background with the presence of noise. Naturally the data set could as well be expanded with more augmentation variants, if that would be of interest.

3.5.Automated Digitization and Summarization of Analog Archives: Comparing Summaries made by GPT-3 and a Human

This section is a summary of the work conducted for the master thesis “Automated Digitization and Summarization of Analog Archives: Comparing Summaries made by GPT-3 and a Human” supervised by AFRY and in collaboration with Uppsala University. The full master thesis is published in DiVA [27].

3.5.1.Introduction

This thesis aimed to create a tool that could assist climate researchers in their fieldwork. Through dialog with researchers at Stockholm’s University a need and interest for automated digitization and summarization of their handwritten notes could be identified. *One of the main challenges for this master thesis work, was that the summarization and text recognition was to be performed in the Swedish language.*

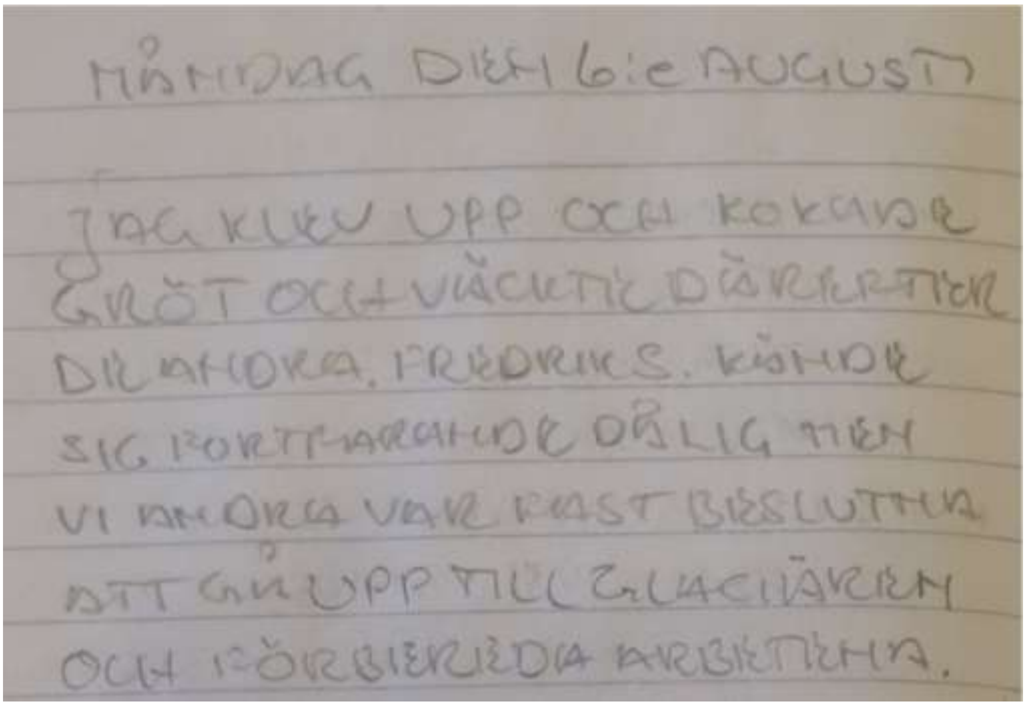
Climate research may require work conducted out in the field and during fieldwork, and many researchers prefer to take handwritten notes which can generate large physical archives. A downside with only physical archives is that the data and knowledge stored here become less available and create a threshold for researchers to use the data since manually digitizing handwritten texts can be very time-consuming.

3.5.2.Tasks and Scope

The data used in this thesis were of images of field journals from Tarfala Research Station, a sample presented in **Figure 12**. For the digitization of the handwritten text, the Google Cloud Vision API [28] was utilized. It was implemented in the program and evaluated for the data used in this thesis. For the summarization, the private algorithm GPT-3 API [29] was used. Two different engines of GPT-3 performed two separate summaries for each data sample and an evaluation of which engine creates the best summaries was investigated with the help of a survey.

To summarize, the thesis aimed to answer the questions:

1. Is it possible to build a program that can digitize handwritten text on a physical object and summarize it?
2. Which GPT-3 engine is recommended for summarizing the data used for this thesis (field journals with handwritten text)?
3. How does the Google Cloud Vision API perform on the data used in this thesis?


MÅNDAG DEN 6:E AUGUSTI

JAG KLEV UPP OCH KOKADE
GRÖT OCH VÄCKTE DÄREFTER
DE ANDRA. FREDRIK S. KÄNDE
SIG FORTFARANDE DÅLIG MEN
VI ANDRA VAR FAST BESLUTNA
ATT GÅ UPP TILL GLACIÄREN
OCH FÖRBEREDA ARBETENA.

Figure 12: A cropped version of a data sample from the data set originating from a field journal from Tarfala Research Station used in this thesis

3.5.3.Purpose and Goals

Based on the results and knowledge from the preliminary study, D6.1, the purpose of this master thesis took shape. There exists a need for a simplified way of digitizing analog data that significantly shortens the time taken to manually digitize a file. A potential need for further processing of text can be identified as a short summary and can give scientists an overview of the main content of a text. Therefore this master thesis investigated how you can digitize and summarize handwritten text on a physical object with the help of AI, create a program researchers can use to digitize and summarize their own data, and evaluate the tools that make up the program. In order to achieve these goals, the work was stated as follows:

- Gain access to GPT-3
- Successfully make API calls on both Google Cloud Vision and GPT-3
- Connect the digitization and summarization to create a pipeline
- Find flaws and strengths in Google Cloud Vision and GPT-3
- Get at least 30 responses on the survey to evaluate the performance GPT-3

3.5.4.Results and Conclusions

A software program was created which could automatically digitize and summarize handwritten texts to save time for researchers. The tool consisted of (1) Google Cloud Vision API used to digitize a photo of handwritten text by using a convolutional neural network (CNN) and (2) the transformer-based algorithm GPT-3 used to summarize the digitized text. The GPT-3 algorithm provided two different engines, Davinci and Curie. The performance of the algorithms was evaluated with a data set consisting of handwritten texts provided by Stockholm's University.

In Figure 13, the summaries generated by the GPT-3 engines, Davinci-Instruct-Beta and Curie-Instruct-Beta are presented. The human-made summaries are also presented.

Note: For brevity, only two examples have been included from the master thesis in this summary.

Original text

Lördag 4 augusti
Morgonen var kall. En isande vind gjorde det riktigt ruggigt när solen gick i moln, vilket den gjorde ganska ofta. Fredrik i kvadrat var som vanligt sist iväg. Vi fick liksom igår vänta en dryg halvtimme. De fem kilometrarna var lättgångna ner till stugan som visade sig vara riktigt trevlig, med kamin och fotogenkök. Alla verkade trötta, men efter ett uppfriskande bad blev stämningen något bättre och vi satt och pratade och läste en stund. Fredrik S började låta lite illa med en rejäl hosta. Storkok av fruktsoppa på kvällen.

Output Davinci-Instruct-Beta	**Output** Curie-Instruct-Beta	**Output** Human
Lördag 4 augusti Morgonen var kall och isande vind gjorde det ruggigt. Solen gick i moln och vi fick vänta en halvtimme på Fredrik i kvadrat. Vi gick 5km ner till stugan som var trevlig. Alla var trötta men efter ett bad blev stämningen bättre. Fredrik S började hosta.	Lördag 4 augusti Morgonen var kall. En isande vind gjorde det ruggigt när solen gick i moln. Fredrik i kvadrat var som vanligt sist iväg. Vi fick vänta en dryg halvtimme. De fem kilometrarna var lättgångna ner till stugan som visade sig vara riktigt trevlig, med kamin och fotogenkök. Alla verkade trötta men efter ett uppfriskande bad blev stämningen något bättre och vi satt och pratade och läste en stund. Fredrik S började låta lite illa med en rejäl hosta. Storkok av fruktsoppa på kvällen.	Lördag 4 augusti Morgonen var kall. En isande vind gjorde det ruggigt. Fredrik i kvadrat var sist. Vi fick vänta en dryg halvtimme. De fem kilometrarna var lättgångna ner till stugan som visade sig vara trevlig. Alla verkade trötta. Efter ett bad blev stämningen bättre. Fredrik S började låta lite illa. Storkok av fruktsoppa på kvällen.

(a) Data sample from 4th of August

Original text

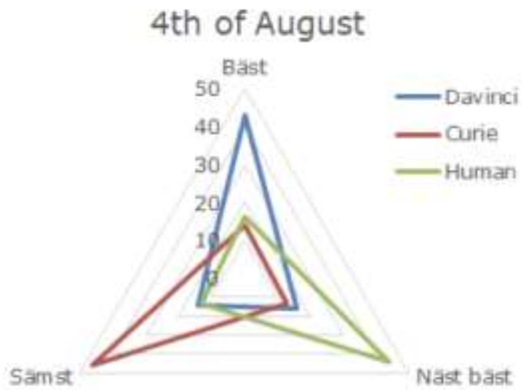
Tisdag 7 augusti
Först mätningar på Vartas till lunch, därefter utflykt till isfallet på Suottas. Sen gick Karsten, Janne, Åsa och Anna tillbaka och fortsatte med arbeten till 00:30.

Output Davinci-Instruct-Beta	**Output** Curie-Instruct-Beta	**Output** Human
Tisdag 7 augusti Vi gjorde mätningar på Vartas till lunch. Sen gick vi till isfallet på Suottas. Karsten, Janne, Åsa och Anna fortsatte arbetet till 00:30.	Tisdag 7 augusti Vi mätte Vartas till lunch och sedan gick vi till isfallet på Suottas. Sen fortsatte vi med arbete till 00:30	Tisdag 7 augusti Mätningar på Vartas vid lunch följt av isfallsutflykt på Suottas. Sen arbetade vi till 00:30.

(b) Data sample from 7th of August

Figure 13: Data samples containing the original text and their corresponding summaries constructed by the GPT-3 engines Davinci-Instruct-Beta, Curie-Instruct-Beta and the author of this thesis

The survey received 73 responses. Out of those people 56,2% were male, 39,7% female, and 4,1% did not want to state their gender identity. 63% were in the age group 21-30, 13,7% were 31-40 years old, 5,5% were 41-50 years old, 9,6% were 51-60 years old. The following graphs visualize how the 73 subjects graded summaries in Figure 14.



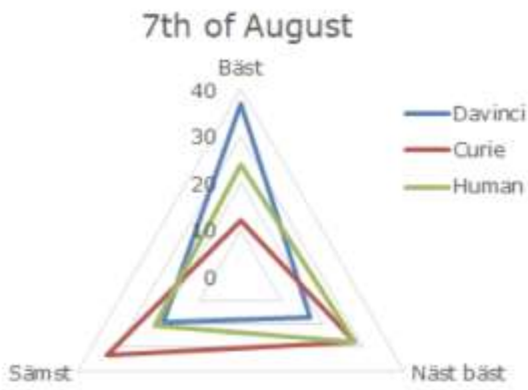


Figure 14: Results from the survey visualizing the distribution of grades for the summaries made by Davinci-Instruct-Beta, Curie-Instruct-Beta and a human

The results indicated that the performance of Google Cloud Vision API was highly correlated to the quality of the image and the way of handwriting. With a unique handwriting follows a poor classification of letters since the algorithm performed badly on shapes that were unfamiliar. A survey was used to evaluate the performance of GPT-3. The survey received 73 responses where the subjects would grade five summaries conducted by a human and the GPT-3 engines Davinci and Curie respectively from the same text. The results from the survey indicated that the performance of the engine Davinci was comparable to the performance of a human while Curie was not a preferable option.

3.6.Computer Vision for Camera Trap Footage: Comparing classification with object detection

This section is a summary of the work conducted for the master thesis "Computer Vision for Camera Trap Footage: Comparing classification with object detection" supervised by AFRY and in collaboration with Uppsala University. The full master thesis is published in DiVA [30].

3.6.1.Introduction

Monitoring wildlife is of great interest to ecologists and is arguably even more important in the Arctic, the region in focus for the research network INTERACT. Camera traps are an effective way to monitor wildlife, where the data are used to study how populations vary over time or how human presence affects the animals.

This master thesis studies how AI and CV can be used together with camera traps to achieve an effective way to monitor populations. The study uses an image data set, containing both humans and animals. The images were taken by camera traps from ECN Cairngorms, a station in the INTERACT network. ECN's pictures feature multiple animal species and humans, but there are also many empty pictures. Examples are shown in **Figure 15**. Sorting and recording information for all the produced images is tedious and time-consuming, Sharp [31] estimates that at ECN, one person processes about one image per minute. This is time that researchers could spend on more complex problems. That is where computer vision and AI, comes in, as a way to quickly process all images, allowing researchers to spend more time on advanced research.

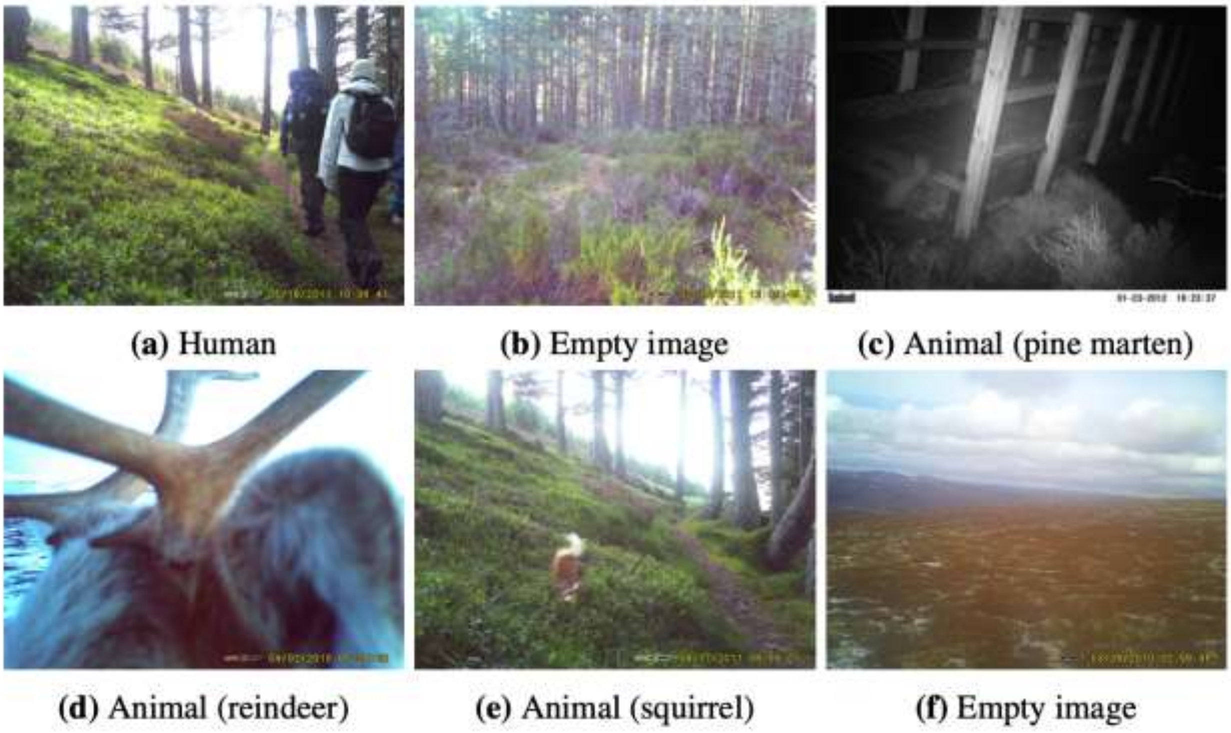


Figure 15: Examples of the pictures making up the data set from the Cairngorms camera traps.

3.6.2.Tasks and Scope

The goal of the project is to classify images into one of three categories: "Empty", "Animal" and "Human".

Developing a model that is useful to the ECN and INTERACT can be divided into two main parts:

1. The model should be general and easy to apply in new settings and on new data sets, in order to maximise its usefulness to the INTERACT researchers, whom will have different data sets and limited previous experience with ML.
2. The choice of method should be technically motivated, for its classification results to be as good as possible and so that researchers using it will know what performance to expect.

To ensure that a model that fits these criteria is found, three different methods were compared:

- DenseNet201 [32], a classification method
- YOLOv3 [33], a detection method
- MegaDetector [18], a pre-trained detector

The choice of which model to use to automate CT image analysis is not obvious and this study aimed to arrive at a recommendation to INTERACT. Beyond developing a useful model, the thesis also aimed to research how well-suited classifiers and object detectors are, compared to each other, when they are used for image recognition on camera trap data.

3.6.3.Results and Conclusions

DenseNet201 and YOLOv3 were both trained on the ECN data set. The pre-trained detector named MegaDetector is developed by Microsoft and is pre-trained on millions of camera trap images from different locations, but none from the Cairngorms.

No sufficient results were achieved with the classifier, DenseNet201, but YOLOv3 performed well on human detection, with an average precision of 0.8 on both training and validation data. The animal detections for YOLOv3 did not reach as high average precision and this was likely because of the smaller amount of training examples.

The best results were achieved by MegaDetector in combination with an added method to determine if the detected animals were dogs, reaching an average precision of 0.85 for animals and 0.99 for humans. This is the method that is recommended for future use, but there is potential to improve all the models and reach even more impressive results.

For the majority of the pictures, the bounding boxes fit very well, and the classification was correct. The model also managed to make detections on many images that seem difficult, see e.g., the latter two images in **Figure 17**. Even though the results were good, we can see that there are approximately 20 % of the animal images that were tough for MegaDetector. Examining the images with undetected animals, provided some insights on what the reason for this can be. Examples of missed detection are seen in Figure 16.

(a) Motion blurry deer image **(b)** Badger behind vegetation **(c)** Fox in strong backlight **(d)** Small squirrel (5x zoom)

Figure 16: Examples of animal images that MegaDetector did not manage to detect.

This project shows that applying CV techniques on CT data is a feasible and effective way to run wildlife monitoring projects and INTERACT (and ecologists in general) should consider implementing the method. It is of interest for stations that want to start monitoring or improve the current monitoring of populations, and as a tool to get a better view of how human activity around stations affect wildlife (although there are privacy measures needed to be taken when humans are pictured). The author estimates that a basic approach with MegaDetector going through CT-images will require less money, occupy less of the researchers' time and be easier to do systematically than the approaches presented in INTERACT's current minimum monitoring programme for fauna [19], where fauna monitoring is said to be ”generally very time-consuming and dependent on many hours of field work.”. Even though the output detections of animals and humans will still need to be reviewed by researchers for advanced analysis, no time will have to be spent on selecting the images of interest.

(a) ECN:s station manager (b) Two detected reindeer. (One un-detected far back.) (c) Close up of human hand (d) Ears of a deer

Figure 17: Examples of correct MegaDetector-detections with blue bounding boxes for humans and red for animals.

3.7.Searching and Recommending Texts Related to Climate Change

This section is a summary of the work conducted for the master thesis “Searching and Recommending Texts Related to Climate Change” supervised by AFRY and in collaboration with Uppsala University. The full master thesis is published in DiVA [34].

3.7.1.Introduction

Users of applications such as Netflix and Spotify have encountered a recommender system. When using e.g., Spotify, the recommended products are songs, artists, and podcasts. The system uses information about the user, such as what they have previously listened to, to recommend music that the user will most likely enjoy.

Aggarwal, in his book about recommender systems [35], claims that the main goal for recommender systems is to increase product sales. This statement is based on the recommended items being products for sale but can be applied to texts as well. We can look at two cases; one with texts behind paywalls and another when the texts are open for everyone. In the first case, relevant texts recommended to the user will hopefully lead to more users buying reports or subscribing to services providing reports and therefore increasing revenue. In the second case, the aim is not to make money but to spread information. If users of the systems have easier access to reports of interest, they are more likely to read the information. It also increases the visibility of similar reports or articles.

A key feature of an information system is to make it easier to browse and find information relevant to the user. This project focuses on the implementation and testing of a way to represent texts about climate change. One desired outcome of this is a snowball effect where people will find interesting information about climate change, allowing a faster spread of knowledge, which in turn might lead to more action being taken in this area.

To make use of recommender systems in the context of items containing texts, NLP is used. NLP can be seen as the transformation of natural language into a language that a computer can understand. It contains, among other things, fetching of the text data, dividing it into words, removing irrelevant words (e.g., prepositions, noise, very rare words), and representing the words as vectors [36].

3.7.2.Tasks and Scope

The goal of the project was to investigate different recommender systems and their usefulness in the context of texts related to climate change research. The project considered the design of an ML system to efficiently search a database of texts related to climate change. The efficient search and navigation of such a database made it easier to find actionable information, detect trends, or derive other useful information. A key feature of such an information retrieval system is the numerical representation of such a text.

Three different approaches were used when implementing a recommender system, in order to see which of them would work best in the setting of texts related to climate change. The different approaches were Bag of Words [37], TF-IDF [38] and Doc2vec [39], where the first two methods do not take the order of words into account, but the last one does. Each method structures the documents as vectors, which are then transformed in different ways within the vector space they exist in.

3.7.3.Results and Conclusions

Google’s Dataset Search was used to find a climate change dataset provided by the University of Illinois at Chicago [40]. Of 30,000+ abstracts, 10,000 were randomly chosen and became the dataset. Figure 18 illustrates a PCA plot of the Doc2Vec representations of the abstract. Before training the system, we needed to pre-process the data. This includes tokenization, removing of stop words and punctuation, POS-tagging, lemmatization, stemming, and lowercasing.

The results show that both Bag of Words, TF-IDF and Doc2vec are performing better when having a longer string. This is probably due to the fact that more words give a more specific description of what is required.

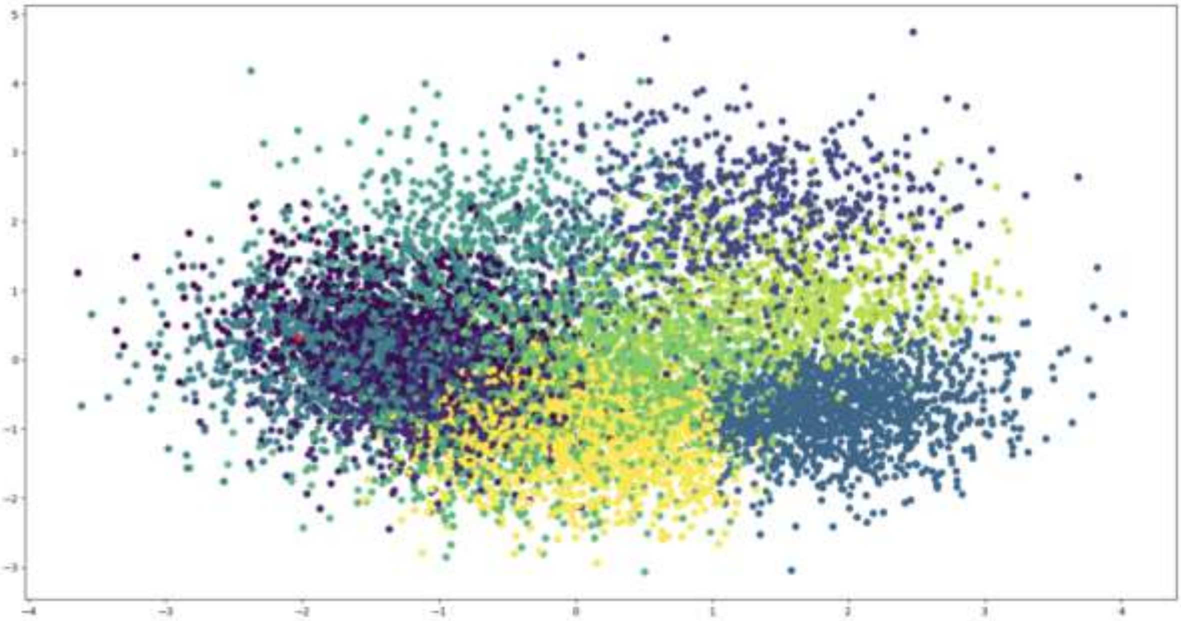


Figure 18: A PCA plot of the Doc2Vec representations of all abstracts in the dataset

Additionally, the reported results indicated two cases: Firstly, we observed that all 3 embeddings outperformed a naive (fixed, expert rule-based) method for retrieving a text. In this case, the query contains part of the text with a small modification, while the result of the query should be the text itself. The Bag-of-Words approach turned out to be best in class for this task. Secondly, we considered the task where the query is a random string, while the desired result is based on a manual comparison of the results. Here we observed that the Doc2Vec approach is best in class. If the random queries become abstract-like, the Bag-of-Words approach were performing almost as well.

3.8.Deep Learning for Iceberg Detection in Satellite Images

This section is a summary of the work conducted for the master thesis "Deep Learning for Iceberg Detection in Satellite Images" supervised by AFRY and in collaboration with Uppsala University. The full master thesis is published in DiVA [41].

3.8.1.Introduction

The application of satellite images for ship and iceberg monitoring is essential in many ways in Arctic waters. Even though the detection of ships and icebergs in images is well established using Geoscience techniques, the discrimination between those two target classes still represents a challenge for operational scenarios.

The remote sensing systems used to detect icebergs are housed on satellites over 600 kilometers above the Earth. The one used to monitor Land and Ocean is the Sentinel-1 satellite. In SAR (Synthetic-aperture radar) images, ships and icebergs typically have a stronger backscatter (the energy reflected back to the radar) response than the surrounding open water and are therefore detectable using adaptive threshold techniques [42]. In general, the surrounding open water will be darker at a higher incidence angle, and thus it is also necessary to consider the radar polarization, which is how the radar transmits and receives the energy. More advanced radars like Sentinel-1, can transmit and receive in both the horizontal and the vertical planes. In this way, a dual-polarization image can be obtained. The data used in this project have two channels: HH (transmit/receive horizontally) and HV (transmit horizontally and receive vertically), which lay an important role in the object characteristics, since objects tend to reflect energy differently.

3.8.2.Tasks and Scope

The main objective of this thesis was to develop and compare three methods for object detection, namely: support vector machines (SVM) [43], convolutional neural networks (CNN) [44] and single shot multibox detector (SSD) [45].

To achieve this, the following objectives should be met:

- Implementing three models (SVM, CNN, SSD) for object detection.
- Comparing these three methods on the SAR (synthetic-aperture radar) imaging dataset from Kaggle (a subsidiary of Google, it is an online community of data scientists and ML practitioners) as a case study, to see how different algorithms perform in iceberg detection which will help transport in polar area and hence to cut-down air pollution.

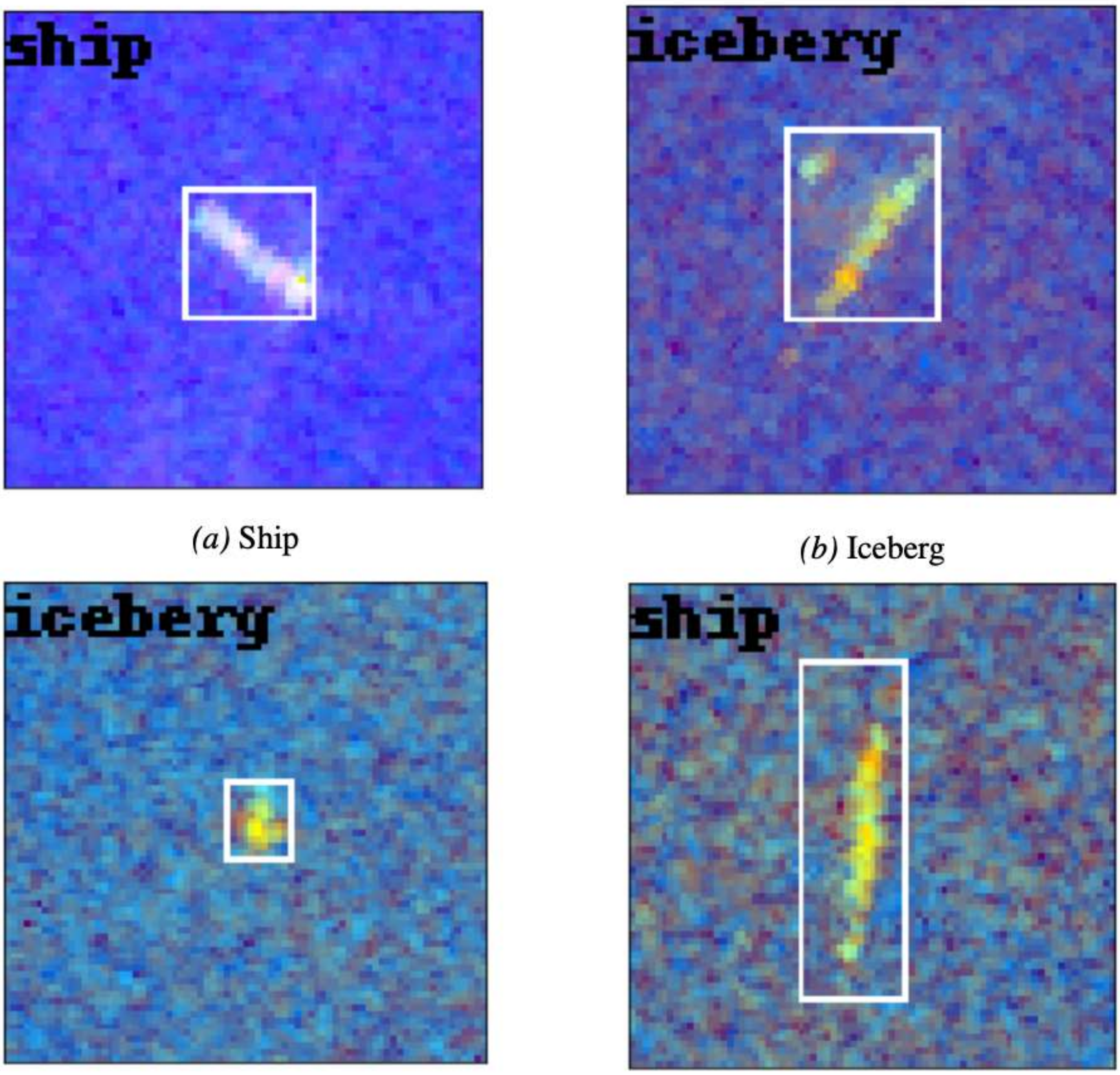
3.8.3.Results and Conclusions

The data used for this study came from the Kaggle "Statoil/C-CORE Iceberg Classifier Challenge" competition [46]. The labels were already provided by human experts and geographic knowledge of the target. Figure 19 illustrates the results of the SSD detection algorithm, yielding a correctly classified ship, a correctly classified iceberg, a misclassified ship and a misclassified iceberg.

Results showed that the CNN model outperformed the SVM model and the SSD model. The performance of the SSD model is second. The SSD has the disadvantage of a longer training time but is able to learn relevant features from the input image, resulting in a good, generalized model and also detect the location of the iceberg in the image. The SVM model has the disadvantage of even longer training time than SSD and lower evaluation scores.

3.9.Collaboration with INTERACT III Research Stations

In addition to the previously mentioned conducted work, this project has also spawned collaborations, and one of the more notable collaborations has been with the Mukhrino field station of Yugra State University, where AFRY has contributed to a discussion piece on the topic of organism classification with the aid of AI and ML [47]. This is a great example of how INTERACT III brings people from different backgrounds together to improve climate research.



(a) Ship

(b) Iceberg

(c) Ship

(d) Iceberg

Figure 19: SSD detection results, (a) Example of a correctly classified ship. (b) Example of a correctly classified iceberg. (c) Example of a misclassified ship. (d) Example of a misclassified iceberg.

4. Future Strategy and Planning

This section reports on the compiled experience gained from WP6 on the future strategy and planning for the area of AI and ML that can be applied in Arctic research. Given the varying nature of the research of the members of and stations associated with INTERACT III, reporting can consequently not be fully applicable to all parties.

4.1.Opportunity Analysis

In short, a multitude of potential applications and use cases can be made for Arctic research.

A vast number of technologies, products, and solutions exist to support development. Some state-of-the-art solutions exist by using APIs, which give limitations when Internet access is not applicable. Other state-of-the-art solutions exist as open-source or can be made available from commercial options. Available and fully operational solutions do exist for Arctic research. Still, to ensure and govern those solutions to be well-designed, -modeled, -developed, and tested for the respective use case and domain, knowledge and expertise in the area are needed.

4.1.1.Potential Use Cases of AI Technology

Results from the pre-study, D6.1, showed that several ideas with large potential on how to utilize AI algorithms were put forth. Examples of those were to assist in finding patterns and making good use of large amounts of data. However, the most commonly mentioned application was to use automatic classification of image data, e.g., removing empty pictures from the data set, classifying animals as well as drawing conclusions concerning anthropogenic influences of wildlife. Detecting invasive species was mentioned in the same context.

Furthermore, AI as an aid in phenology monitoring, both tracking cycles and e.g., automatically counting flowers was brought up. There was also an interest in identifying bird songs and in connection help counting birds. Further, collecting, merging, and analysing data from diverse data sets and drawing inferences from these, e.g., how weather data affect reindeer grazing quality. Similarly, a potential AI application was to flag when monitoring data were suspected of being incorrect.

With regards to social studies, it was expressed that it would be a great help if AI and ML could aid in extracting relevant answers from surveys; also, in several languages, using automatic translations.

Drones and how to use them in connection with AI were also a popular topic; for instance, using automatic recognition in quantifying riverbank erosion. In one case, a researcher was investigating what technical specifications of drones and camera equipment would be needed to use DL applied on glaciers or looking at hot spots regarding gas emissions. Similarly, SAR data was brought up for detection of geo-events, as well as using AI with remote sensors, for automatic image analysis on-site.

4.1.2.Open-Source Solutions

Within the fields of AI, experts rely on open-source tools, packages, models, and programming languages. To leverage these technologies, knowledge of using the technologies and domain expertise is required. In the field of AI, framing or defining the problem, as well as obtaining qualitative training data for that problem are the two most common tasks.

Understanding how to frame or design the problem requires the expertise of the technologies, with a minimum consolidation of domain knowledge. This task can roughly be seen as adapting, framing, and translating data for enabling a computer to understand and a model of analysis to answer a defined question.

To prepare a qualitative and robust model, there is a need for qualitative, quantitative, and representative data for what the model is to be exploited. Tools, such as Microsoft's Power BI, exist to better understand the data and what questions can be answered with that data. Gathering qualitative and quantitative amounts of training data is often necessary in the case of ad-hoc development, since appropriate and representative training data for new use cases in new domains and applications rarely exist.

In the development process when the AI application has a defined problem, and training data exist in an analysis ready format, the selection of models can be implemented using open-source libraries or tools. Examples of those are Keras, Tensorflow, PyTorch and Scikit-Learn. When informed choices have been made, empirical testing and evaluation are used for validating the model selection and adaptation.

4.1.3.Commercial Solutions

Many commercial providers and suppliers exist in the field of AI. Today, it is possible to upload training data to any of the large cloud providers (like Microsoft Azure, Amazon Web Services, or the Google Compute Engine) for them to automatically suggest use cases, selection of models, and applications. It is also possible to order on premise solutions, computation or storage, end-to-end implementations etcetera. The different ways of using large as well as smaller providers are endless and depend on what is to be created, on requirements mapped to different technical trade-offs, and on the knowledge of the technologies which are to be used.

4.1.4.Potential Methodologies for AI Applications

During the work with WP6, many potential use cases of AI for Arctic research have been presented. Without emphasizing too much on a specific use case, this section will give examples of how AI applications can be developed in different ways. Differing development methodologies exist and can be seen to depend on the implementation level and strategy, as well as the technical use case. In the following section, three examples will be presented.

A methodology of creating an initial proof-of-concept for an AI use case by a master thesis student can be outlined in the following steps:

1. List potential resource consuming or manual tasks for personnel, suitable for an automatic workflow, which is not too time critical, and does not need a live data stream.
2. Consult an expert within AI applications, discuss the potential technical feasibility of making each task a potential AI technological use case, and select a handful of use cases.

3. Hire master thesis students from the IT, data engineering, data science, or AI domain to investigate and set up a PoC.
4. Let them present their master thesis work with a focus on learning outcomes and future development.
5. Evaluate what went well, what could be done better, and if this PoC is something that would assist Arctic research.

A methodology for developing an intermediate and narrow AI application with a defined use case which have concrete value for Arctic research, can be outlined in the following steps:

1. Map out the existing data sources available, as deemed necessary for the AI model, with metadata.
2. Let AI experts consolidate on the design and tasks of the application.
3. Let the AI experts suggest functionalities together with a graphical user interface to adopt user feedback on the AI results.
4. Map out requirements and core needs with small iterations of graphical interface and consolidation with representative potential users.
5. Set up an offline and cloned dataset containing the data needed to demonstrate the technical viability of the application.
6. Grant resources to develop and implement a graphical interface, functionalities, and application.
7. Test functionality and demonstrate the application, as well as foreseen outcomes for implementation of such application.

A methodology for implementing real time AI analysis of specific occurrences with weekly reports of results, without connection to the Internet and sensors, can be outlined in the following steps:

1. Grant resources and access to data, sources of computation, and storage (such as an applicable computer) and sensors for development.
2. Define what a specific occurrence is, given the sensors' data that already exist.
3. Define how the information of a specific occurrence is best represented and presented to the key personnel.
4. Define how to technically update what is to be defined as a specific occurrence.
5. Define the system requirements.
6. Consult AI experts in putting the solution together with the consolidation of key personnel.
7. Test the solution.
8. Implement the solution.

9. Update and monitor the functionality.
10. Life cycle management.

4.2. Future Needs of AI Technology

As was outlined in Section 3.3 in the pre-study, D6.1, funding is an issue for many stations, and if AI can help save man hours giving increased resources for more advanced research, many researchers would be eager to incorporate AI and ML in their workflow.

With technical development and more sensors, more data is foreseeable to be created at a faster pace, leading to increased accumulation of even more data and the need for automatic data analysis. Monotonous work, for instance manually classifying animals in photographs, counting birds, plants, or samples is very time-consuming and tedious, and not a very good use of already lacking resources. These resources could more efficiently be utilized by performing analysis and research.

During the discussions of the workshops held by WP6, it became apparent that many researchers were new to AI, and while curious, had little knowledge about it and were unsure how to practically proceed. From the interviews of the pre-study, the general experience is that interviewees have a positive view of AI while being unclear on particulars of what it can and cannot do, as well as lacking skills in how to practically proceed with AI. One researcher also mentioned that not everyone had enough experience in mathematics to grasp AI, while others stated that they in general were unsure of how to approach AI practically. Results from the pre-study showed that a way to utilize AI and ML without having to program would be preferred.

4.3. Future Goals

AFRY suggest defining future goals for increasing the use of AI in Arctic research. These are presented without a proposed order but as separate entities.

Firstly, it is important to understand where the data-driven maturity is at in the organization. This is best done by assessment from a partner looking from an outside perspective since it can look different from within different places in an organization.

Secondly, listing manual and resource consuming tasks, a selection of viable and feasible use cases of AI can be chosen to assist in the process before developing initial proof-of-concepts. This can assist in raising further interdisciplinary knowledge of AI technology in the respective research domain and lead to later functional implementations.

Thirdly, increase the structure of data, and level of quality of metadata and map out where the access to that data can be found. This work is something that INTERACT III has as one of its goals and its further increase will make it possible for easier implementations of AI solutions.

Lastly, adopt a more data driven operations and/or maintenance approach to Arctic research. Saving time and resources for different processes of the research can lead to more increased time and resources for advanced research.

4.4.Identified Initial Steps

AI is well suited for automizing manual tasks. Currently, AI technology is mature to assist in narrow applications. AFRY recommends stations list their most manual and repetitive tasks that are estimated to require an unnecessarily large number of resources. Examples of this can be found in the results of the pre-study [15]. Fieldwork and subsequent analysis and documentation, performed by Arctic researchers, are in many cases time consuming, with tasks such as:

1. Looking at photographs and identifying and classifying plants in various stages.
2. Identifying and classifying animals and their species, as well as counting them.
3. Listening to bird songs and identifying the bird species.

After listing, a selection of viable and feasible tasks could be developed as proof-of-concept projects, by students or personnel knowledgeable in or eager to learn more about the technologies. Other alternative approaches to this are also:

1. Simplifying the application of ML by providing “ML as a service”, i.e., providing a simple method for accessing, using, and managing ML functionality, perhaps as a cloud enabled service.
2. Providing support to researchers in applying ML to their data.

To further increase general awareness, understanding, potential use-cases, and applications in AI solutions and technology, the recommended action is to have introductory education of key personnel.

Increased structure for data and metadata can assist in increasing the opportunity for enabling AI to assist in Arctic research. One example, highlighting issues with data, is when language and information from a text data set are in other languages than English. Natural language models today that are considered state-of-the-art are trained in the English language, which poses a problem for leveraging the technology for other languages in an easy manner. Another example highlighting the issues with poor quality in metadata can be when there are no descriptions or information of what the data set contains, where it is from and what it represents. In that case, it is difficult for new observers or users to understand what can be done with it.

AFRY also recommend adding competence from the areas of scientific computing or IT. This can be added in different ways, with processes such as:

- Educating researchers with a focus on practical use of the technology, tools, and libraries.
- Adding competence from the field of technology and the opportunity to consolidate the field of research.

4.5.Identified Challenges

A common theme that emerged during the interviews of the pre-study was obstacles encountered concerning monitoring and research. The most pressing issue was that of funding. No matter the importance, longevity, or even prestige of the research station, funding always crept up as a major hindrance. Funding was mentioned as a constant struggle, even in cases when the research station in question would be used as a promotion for climate change research. Moreover, concerning funding, another factor that was perceived as a potential problem by the informants was that their work was steered towards the desired outcome of the donor.

There was the issue of collecting data, or sensitive instruments needing monitoring and calibration, in many cases, in a high-risk environment and usually a large land area. It was also mentioned that there were sometimes issues with the power supply.

At the location of one station, there was a ban on the use of Wi-Fi and Bluetooth, due to there being a radio telescope sensitive to such frequencies. This makes it more difficult to conduct science, and disables, among other things, the use of drones. It is possible to apply for permission, but that is often a futile process.

The results from the pre-study and workshop showed that added competence is needed. Python is currently the most popular programming language used for AI and ML. The main reason for this is that many Python libraries are easy to use, while still being powerful. However, not everyone can, or have the time or resources to learn how to program Python.

4.6.Suggested Roadmap

In the AI strategy, it is vital to identify the technical and non-technical areas suitable for AI implementation. It is essential to identify areas of quality improvement, technical assistance in decision making, areas where process connections are necessary for full AI implementation, where resource efficiency can be improved, where staffing is under-established or deteriorating etcetera.

In all cases, it is vital to identify which process objects or databases must be accessible for each AI implementation.

Other examples that can assist in increasing the role of AI in Arctic research are:

- More connectivity with sensors and hubs that will collect data in a structured format
- More identified narrow applications with defined tasks
- Increased competence in the field of AI and automatic data analysis
- A look into data driven operations and/or maintenance of the stations, leading to more resources for advanced research

5. Summary and Conclusions

5.1.Summary

Previously undertaken work in accordance with the deliverables for WP6 has consisted of:

1. Pre-study on inquires and needs from identified station managers and researchers, to identify possible datasets and type of questions to be answered
2. Workshop with demonstration on technology available today and expected in the future in the area of ML and AI technology
3. Using ML on example data to make specific algorithms and methods available and demonstrating the outcome
4. Master theses conducted pertaining to AI and ML, applied to Arctic research for INTERACT III, in collaboration with Uppsala University:
    a. "Image Augmentation to Create Lower Quality Images for Training a YOLOv4 Object Detection Model" [12]
    b. "Deep Learning for Iceberg Detection in Satellite Images" [41]
    c. "Searching and Recommending Texts Related to Climate Change" [34]
    d. "Computer Vision for Camera Trap Footage: Comparing Classification with Object Detection" [30]
    e. "Automated Digitization and Summarization of Analog Archives: Comparing Summaries Made by GPT-3 and a Human" [27]

In addition, experience gained from WP6 with regards to future strategy and planning for the area of AI and ML in Arctic research has been compiled. One of the focus areas has been opportunity analysis concerning potential use cases of AI technology both with regards to open-source solutions as well as commercial solutions. Moreover, methodology and guidance on how to develop general awareness, understanding and use-cases in AI solutions and technology, have been outlined.

5.2.Conclusions

The Arctic is a very important region to get an in depth understanding of climate changes. It is not only a region, but also a physical, biological, chemical, and climatological system. Valuable data are collected by dedicated researchers involved in INTERACT III. While not all collected data thus far have a purpose, analyses of these data in the future may provide the missing link to understanding key concepts and bringing crucial insights to climate change research.

The purpose of this document has been to put into context previously conducted work. This concerns both previous deliverables of WP6, as well as master thesis work. In addition, an integral part of this document has been to report on future strategies and planning for the area of AI and ML to be applied in Arctic Research.

Workshops were held as reported in D6.2 [16]. These aimed to give insights and a foundation of AI and ML, for researchers to come up with their own application areas. With this as a background, a pre-study was undertaken in D6.1 [15], with the purpose of investigating what inquiries and needs researchers and station managers have, what research is conducted and what data are collected at the stations, as well as what obstacles are encountered by researchers. The results of the study showed that several research groups are keenly interested in applying AI and ML in their research, but do not know how to do so. The perception appeared to be that the main use of AI and ML is to support in decreasing the amount of tedious and error-prone field work and associated analysis and documentation. The contents of this report served as input to the deliverables D6.3 [17], and this document, D6.4, as well as informed the subject matter of several of the master theses undertaken.

D6.3 provided a demonstration and a presentation of pilot projects in which AI and ML techniques were applied on researchers' data. The demonstration showed that many exciting possibilities exist for INTERACT researchers to use AI tools to analyze data and create new insights, with or without coding. These possibilities should not be overlooked, and it is highly recommended for smaller research projects to start utilizing the demonstrated techniques.

With regards to the master thesis work conducted under INTERACT III, the insights brought from "Image Augmentation to Create Lower Quality Images for Training a YOLOv4 Object Detection Model" were that it is important to understand what kind of data have potential to be used in AI. Furthermore, it was investigated how the cutting-edge YOLOv4 object detection model would react to various disturbances in the data set. A model was successfully trained and a correlation between worse performance and presence of noise was detected, but changes in saturation and altered color levels seemed to have less impact than expected. Reducing noise in gathered data is seemingly of greater importance than enhancing images with lacking color levels.

In "Computer Vision for Camera Trap Footage: Comparing Classification with Object Detection", it was investigated how AI and CV can be used together with camera traps to achieve an effective way to monitor populations. The study uses an image data set, containing both humans and animals. The images were taken by camera traps from ECN Cairngorms, a station in the INTERACT network. The goal of the project was to classify images into one of three categories: "Empty", "Animal" and "Human". Three different methods were compared, a DenseNet201 classifier, a YOLOv3 object detector, and the pre-trained MegaDetector, developed by Microsoft. The best results were achieved by MegaDetector in combination with an added method to determine if the detected animals were dogs, reaching an average precision of 0.85 for animals and 0.99 for humans. This is the method that is recommended for future use, but there is potential to improve all the models and reach more impressive results.

The application of satellite images for ship and iceberg monitoring is essential in many ways in Arctic waters. Even though the detection of ships and icebergs in images is well established using Geoscience techniques, the discrimination between those two target classes still represents a challenge for operational scenarios.

"Deep Learning for Iceberg Detection in Satellite Images" proposed the application of support vector machines (SVM), convolutional neural networks (CNN), and the SingleShot Detector (SSD) for ship-iceberg detection in satellite images. The CNN model was compared with SVM and SSD, and the final results indicated not only a superior classification performance of the proposed methods but also the object detection results from SSD.

With regards to "Automated Digitization and Summarization of Analog Archives: Comparing Summaries Made by GPT-3 and a Human" a software program was created which could automatically digitize and summarize handwritten texts to save time for researchers. The tool consists of (1) Google Cloud Vision API used to digitize a photo of handwritten text by using a convolutional neural network (CNN) and (2) the transformer-based algorithm GPT-3 used to summarize digitized text.

The GPT-3 algorithm provided two different engines, Davinci and Curie. The performance of the algorithms was evaluated with a data set consisting of handwritten texts provided by Stockholm's University. The results indicated that the performance of Google Cloud Vision API was highly correlated to the quality of the image and the way of handwriting. With a unique handwriting follows a poor classification of letters since the algorithm performed badly on shapes that were unfamiliar. A survey was used to evaluate the performance of GPT-3. The survey got 73 responses where the subjects would grade five summaries conducted by a human and the GPT-3 engines Davinci and Curie respectively from the same text. The results from the survey indicated that the performance of the engine Davinci was comparable to the performance of a human while Curie was not a preferable option.

"Searching and Recommending Texts Related to Climate Change" considered the design of an ML system to efficiently search a database of texts related to climate change. This project implemented and compared three different ways to represent a text in a vector space. Specifically, Bag-of-Words, Term Frequency - Inverse Document Frequency, and Doc2Vec were contrasted in this context.

The reported results indicated two cases: firstly, it was observed that all 3 embeddings outperformed a naive (fixed, expert rule-based) method for retrieving a text. In this case, the query contains part of the text with a small modification, while the result of the query should be the text itself. The Bag-of-Words approach turned out to be best in class for this task. Secondly, we considered the task where the query is a random string, while the desired result is based on a manual comparison of the results. Here we observed that the Doc2Vec approach was best in class. If the random queries became abstract-alike, the Bag-of-Words approach was performing almost as well.

Looking into future strategies for AI and ML; in order to further enable opportunities to adopt a more data driven approach to research or assisting the work of Arctic research, there is a need to understand the current state where this technology is to be implemented, in terms of data driven matureness. The current state of matureness should be assessed with humans, IT, and processes in mind, and can assist in identifying the terms and requisites for increased use of the technology.

The INTERACT III project is a great way to, among other things, increase the collaboration with actors investigating several areas with a data driven perspective. Further collaboration in the field of AI and ML can result in shared technical applications assisting the work.

Added competence or increased allocation of resources like data engineering/science and IT competence, can result in an increased matureness and opportunity of adopting a more data driven approach. Domain knowledge is required to set up applications or solutions, where experts can consolidate with researchers.

Developing a new AI solution or application for a new domain or use case, needs to be framed or translated into a problem formulation interpretable by an AI model for ensuring proper implementation, which is a task that can require experience and knowledge of the technologies. Often, intermediate levels of implementations can be more fruitful than more advanced levels that require more resources to properly set up for use. Examples of this are that existing AI or CV applications can assist in monitoring analogue indicators, sensors, dashboards, or critical components for assisting work.

In the same fashion, as a human can observe from their eyesight, sensors can be used to monitor, reducing resource consuming tasks of manual evaluation, continual observation, or do check-ups. Automatically knowing exactly what has happened is not always essential. Knowing where to look can often save many resources and decrease human mistakes, before being further examined by humans.

AI applications give us the possibility to analyse a larger amount of data than humans. This can be done with AI models of different complexities, with varying approaches - with examples ranging from easily interpretable expert systems to complex deep neural networks. Even though there are different complexities or abstractions of solutions, mature AI technology is found for narrow applications or ANI, and a good start is to rather use AI and ML with a less advanced algorithm, which is easier to use, than an advanced "killer" algorithm.

Bibliography

[1] National Snow and Ice Data Center, "All About Arctic Climatology and Meteorology," 2020. https://nsidc.org/cryosphere/arctic-meteorology/index.html (accessed Aug. 17, 2022).
[2] IPCC, "IPCC Fifth Assessment Synthesis Report-Climate Change 2014 Synthesis Report," *IPCC Fifth Assess. Synth. Report-Climate Chang. 2014 Synth. Rep.*, 2014.
[3] S. J. Hassol, *Impacts of a Warming Arctic*. Cambridge University Press, 2004.
[4] J. D. Ford, "Indigenous health and climate change," *Am. J. Public Health*, vol. 102, no. 7, pp. 1260–1266, 2012, doi: 10.2105/AJPH.2012.300752.
[5] M. Chui, "Artificial intelligence the next digital frontier," *McKinsey Co. Glob. Inst.*, vol. 47, pp. 3–6, 2017.
[6] A. Ng, "AI for everyone. Coursera.," 2019. https://www.coursera.org/lecture/ai-for-everyone/week-1-introduction-SRwLN (accessed Aug. 17, 2022).
[7] "Types of Machine Learning Algorithms You Should Know | by David Fumo | Towards Data Science." https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861 (accessed Aug. 17, 2022).
[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[9] "What Is Deep Learning? | How It Works, Techniques & Applications - MATLAB & Simulink." https://se.mathworks.com/discovery/deep-learning.html (accessed Feb. 13, 2021).
[10] R. Rojas, *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.
[11] D. C. C. Cires¸an, U. Meier, J. Masci, L. M. Gambardella, and J. ¨ Urgen Schmidhuber, "Flexible, High Performance Convolutional Neural Networks for Image Classification."
[12] T. Melcherson, "Image Augmentation to Create Lower Quality Images for Training a YOLOv4 Object Detection Model," Signals and Systems, Department of Electrical Engineering, Technology, Disciplinary Domain of Science and Technology, Uppsala University, 2020.
[13] S. Li, "Understanding Word2vec Embedding in Practice |Towards Data Science," 2019. https://towardsdatascience.com/understanding-word2vec-embedding-in-practice-3e9b8985953 (accessed Aug. 17, 2022).
[14] J. McCurry, "South Korean AI chatbot pulled from Facebook after hate speech towards minorities," *The Guardian*, 2021. https://www.theguardian.com/world/2021/jan/14/time-to-properly-socialise-hate-speech-ai-chatbot-pulled-from-facebook (accessed Aug. 17, 2022).
[15] M. Erman, T. Skarstedt, and M. Skogsmo, "D6.1 - A Preliminary Study of Inquiries and Needs from Station Managers and Researchers," *INTERACT III*, 2021, Accessed: Aug. 17, 2022. [Online]. Available: https://eu-interact.org/app/uploads/2021/02/D6.1-1.pdf.
[16] M. Erman, M. Skogsmo, R. Cubo, and T. Skarstedt, "D6.2 - Workshop Report on Artificial Intelligence and Machine Learning for Arctic Research," *INTERACT III*, 2021, Accessed: Aug. 17, 2022. [Online]. Available: https://eu-interact.org/app/uploads/2021/02/D6.2-2.pdf.
[17] M. Erman, M. Skogsmo, C. Sundström, and F. Örn, "D6.3 - Demonstration Report on Artificial Intelligence and Machine Learning," *INTERACT III*, 2022, Accessed: Aug. 17, 2022. [Online]. Available: https://eu-interact.org/app/uploads/2021/02/D6.3-1.pdf.
[18] Microsoft, "Megadetector - microsoft/CameraTraps: Tools for training and running detectors and classifiers for wildlife images collected from motion-triggered cameras."

https://github.com/microsoft/CameraTraps (accessed Aug. 17, 2022).
[19] Boone M. E. and Basille M., "Using iNaturalist to Contribute Your Nature Observations to Science." https://www.gbif.org/ (accessed Aug. 17, 2022).
[20] "Welcome To Colaboratory - Colaboratory." https://colab.research.google.com/ (accessed Aug. 17, 2022).
[21] "Microsoft Azure." https://azure.microsoft.com/ (accessed Aug. 17, 2022).
[22] Google, "Google Open Images Dataset." https://storage.googleapis.com/openimages/web/index.html (accessed Aug. 14, 2022).
[23] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv Prepr. arXiv2004.10934*, 2020.
[24] I.E. Commission, "Multimedia systems and equipment-colour measurement and management-part 2-1: Colour management-default rgb colour space-srgb," *IEC 61966-2- 1*, 1999. https://webstore.iec.ch/publication/6169 (accessed Aug. 15, 2022).
[25] A. K. Boyat and B. K. Joshi, "A review paper: noise models in digital image processing," *arXiv Prepr. arXiv1505.03489*, 2015.
[26] T. S. Aiba and S. S. Stevens, "Relation of brightness to duration and luminance under light- and dark-adaptation," *Vision Res.*, vol. 4, no. 7–8, pp. 391–401, 1964.
[27] M. Linderholm, "Automated Digitization and Summarization of Analog Archives : Comparing summaries made by GPT-3 and a human," Signals and Systems, Department of Electrical Engineering, Technology, Disciplinary Domain of Science and Technology, Uppsala University, 2022.
[28] "Vision AI | Derive Image Insights via ML | Cloud Vision API | Google Cloud." https://cloud.google.com/vision?hl=sv (accessed Aug. 17, 2022).
[29] "GPT-3 Powers the Next Generation of Apps." https://openai.com/blog/gpt-3-apps/ (accessed Aug. 17, 2022).
[30] F. Örn, "Computer Vision for Camera Trap Footage : Comparing classification with object detection," Division of Visual Information and Interaction, Department of Information Technology, Mathematics and Computer Science, Disciplinary Domain of Science and Technology, Uppsala University, 2021.
[31] S. Sharp, "Evidence from Camera Trap Data on the Relationship between Human and Animal Presence in the Allt a' Mharcaidh ECN Site," *Univ. Highl. Islands*, 2019.
[32] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
[33] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv Prepr. arXiv1804.02767*, 2018.
[34] K. Gjöthlén, "Searching and Recommending Texts Related to Climate Change," Department of Information Technology, Mathematics and Computer Science, Disciplinary Domain of Science and Technology, Uppsala University, 2021.
[35] C. C. Aggarwal, *Recommender systems*, vol. 1. Springer, 2016.
[36] T. Beysolow II, *Applied Natural Language Processing with Python: Implementing Machine Learning and Deep Learning Algorithms for Natural Language Processing*. Apress, 2018.
[37] Y. Zhang, R. Jin, and Z.-H. Zhou, "Understanding bag-of-words model: a statistical framework," *Int. J. Mach. Learn. Cybern.*, vol. 1, no. 1, pp. 43–52, 2010.
[38] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to information retrieval*, vol. 39. Cambridge University Press Cambridge, 2008.
[39] "models.doc2vec – Doc2vec paragraph embeddings — gensim."

https://radimrehurek.com/gensim/models/doc2vec.html (accessed Aug. 17, 2022).

[40] "T. Complex and S. U. N. Laboratory." https://csun.uic.edu/datasets.html.

[41] S. Dong, "Deep Learning for Iceberg Detection in Satellite Images," Department of Information Technology, Mathematics and Computer Science, Disciplinary Domain of Science and Technology, Uppsala University, 2021.

[42] C. Bentes, A. Frost, D. Velotto, and B. Tings, "Ship-Iceberg discrimination with convolutional neural networks in high resolution SAR images," 2016.

[43] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[44] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, 2018.

[45] W. Liu *et al.*, "Ssd: Single shot multibox detector," in *European conference on computer vision*, 2016, pp. 21–37.

[46] "Statoil/C-CORE Iceberg Classifier Challenge | Kaggle." https://www.kaggle.com/c/statoil-iceberg-classifier-challenge (accessed Aug. 17, 2022).

[47] F. N., A. D. V., B. Y. M., V. V. Bilous, D. A. Bochkov, and Bolshakov Sergey Yu, "Crowdsourcing Fungal Biodiversity: Revision of INaturalist Observations in NorthWestern Siberia," *Nat. Conserv. Res.*, 2022.

[48] "Gensim: Topic modelling for humans." https://radimrehurek.com/gensim/ (accessed Feb. 08, 2021).