KAUST Academy



SECTION 3

ARTIFICIAL INTELLIGENCE.



academy.kaust.edu.sa

40

3D PRINTING FAILURE DETECTION AND ANALYSIS

AHMAD ALQARNI

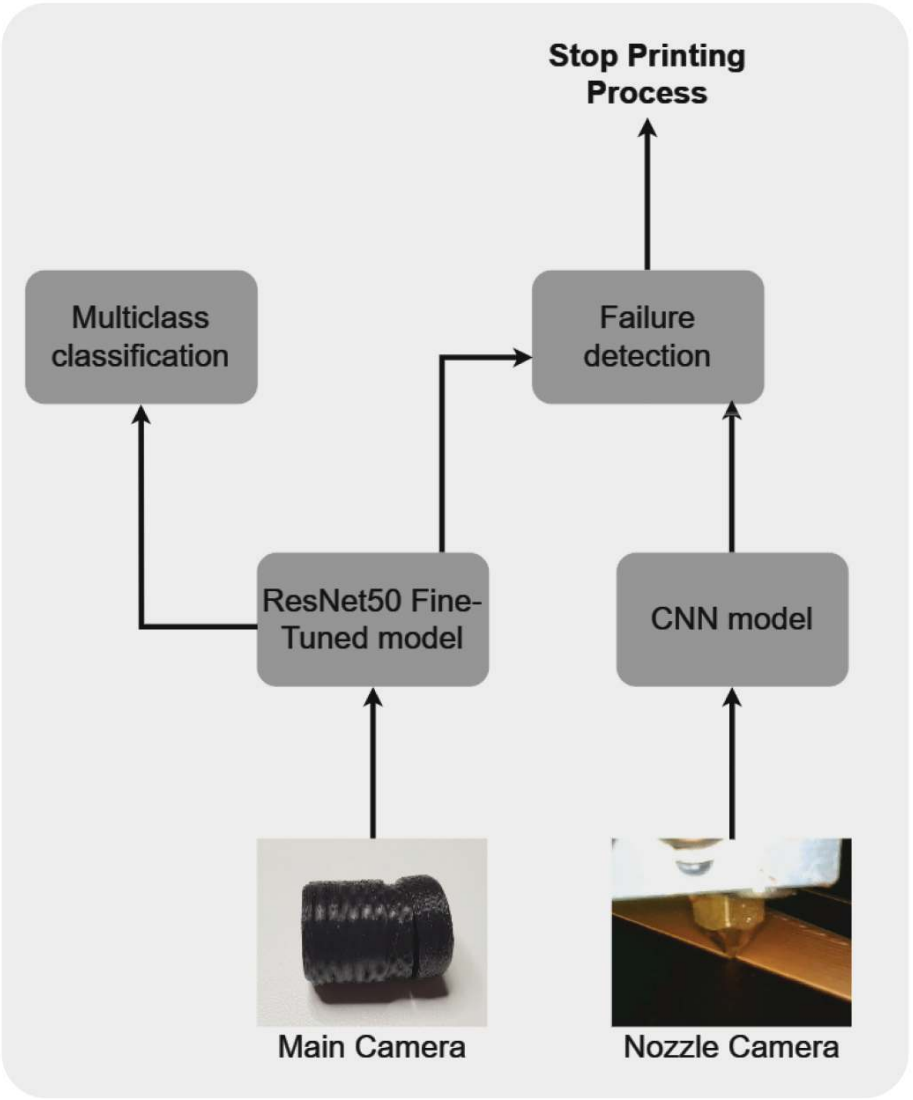
MOHAMMED ALQURAINI



ZIYAD ALJADHAI

AL ABDUL WAHAB

Our deep learning model monitors 3D printing in real-time by analyzing images captured at various intervals from a camera mounted on the 3D printer. The system detects issues such as under-extrusion or spaghetti defects, allowing for early problem identification. When a defect is detected, the model sends a signal to a Raspberry Pi 4 device to halt the printing process, minimizing wasted time and materials.


Stop Printing Process
Multiclass classification
Failure detection
ResNet50 Fine-Tuned model
CNN model
Main Camera
Nozzle Camera

41

ARABIC NLP: GENERATING MULTIPLE CHOICE QUESTIONS OF SAUD SCHOOLBOOKS

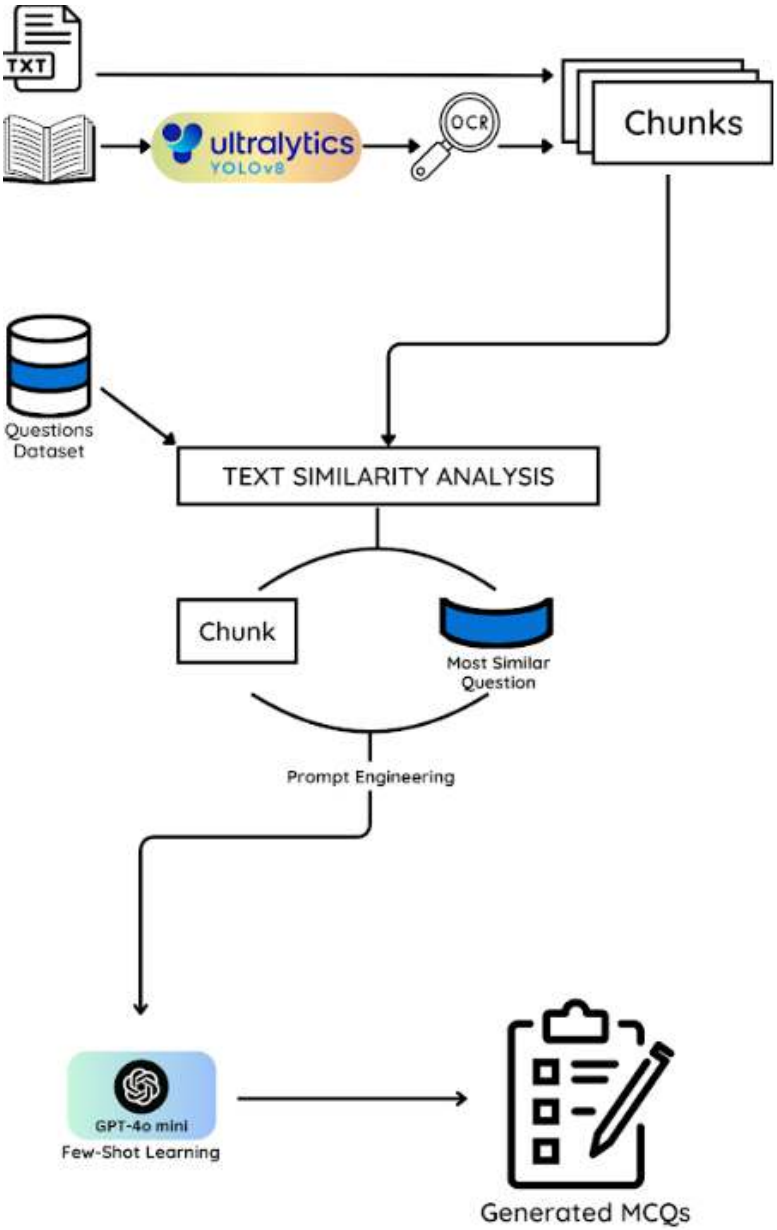

YAZAN YAHYA ALSHUAIBI


MOHANAD AL DAKHEEL


ZYAD RASHED ALZHRANI


MOAYAD ALGHAMDI

Our project aims to advance Arabic Natural Language Processing (NLP) by utilizing Retrieval-Augmented Generation (RAG) and Large Language Models (LLMs) to generate Multiple-Choice Questions (MCQs) from multiple data sources, primarily Saudi schoolbooks. By ensuring that these questions are both contextually relevant and accurate, our approach supports the broader goal of modernizing education, aligning with Saudi Arabia's vision for the future.


TXT
ultralytics
YOLOv8
OCR
Chunks
Questions Dataset
TEXT SIMILARITY ANALYSIS
Chunk
Most Similar Question
Prompt Engineering
GPT-4o mini
Few-Shot Learning
Generated MCQs

DEEP LEARNING-BASED FAULT DETECTION SYSTEM FOR STRUCTURAL DEFECTS IN NANO-QUADCOPTERS

SHAHAD JEZA ALAMRI

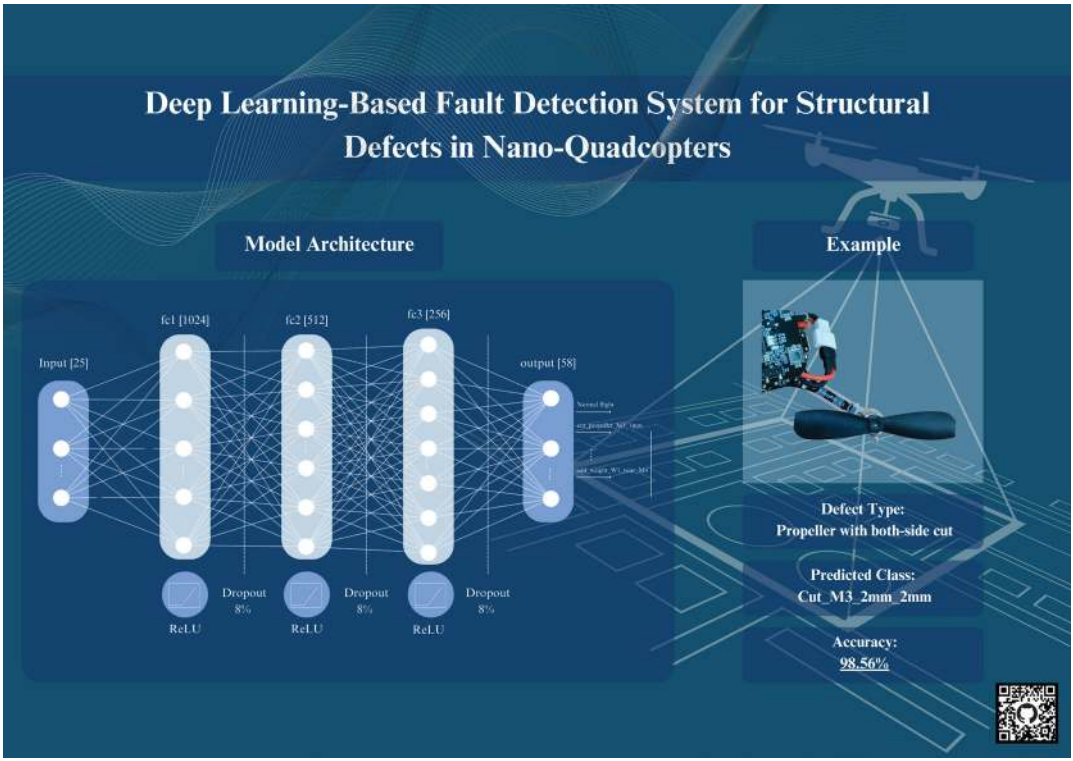
REMA YASIR ALBARAKATI

RAWAN YASIR ALBARAKATI

SALHA NASSER ALYAMI

SHATHA TURKI ALSAIDY

Detecting faults in nano-quadcopters is critical to ensuring safety and performance. This study proposes a deep learning-based system for detecting structural defects in real time. Utilizing advanced feature engineering and dimensionality reduction techniques, the proposed model offers a reliable and efficient solution for ensuring the integrity and functionality of nano-quadcopters in various applications with high accuracy and low inference time.


Deep Learning-Based Fault Detection System for Structural Defects in Nano-Quadcopters
Model Architecture
Example
Input [25]
fc1 [1024]
fc2 [512]
fc3 [256]
output [58]
Dropout 8%
Dropout 8%
Dropout 8%
ReLU
ReLU
ReLU
Defect Type:
Propeller with both-side cut
Predicted Class:
Cut_M3_2mm_2mm
Accuracy:
98.56%

CHARTWISE: CASE-SPECIFIC MULTIMODAL TOOL FOR ENHANCED DATA-DRIVEN ANALYTICS


SAMEER ALSABEA


ABDULLAH JAN

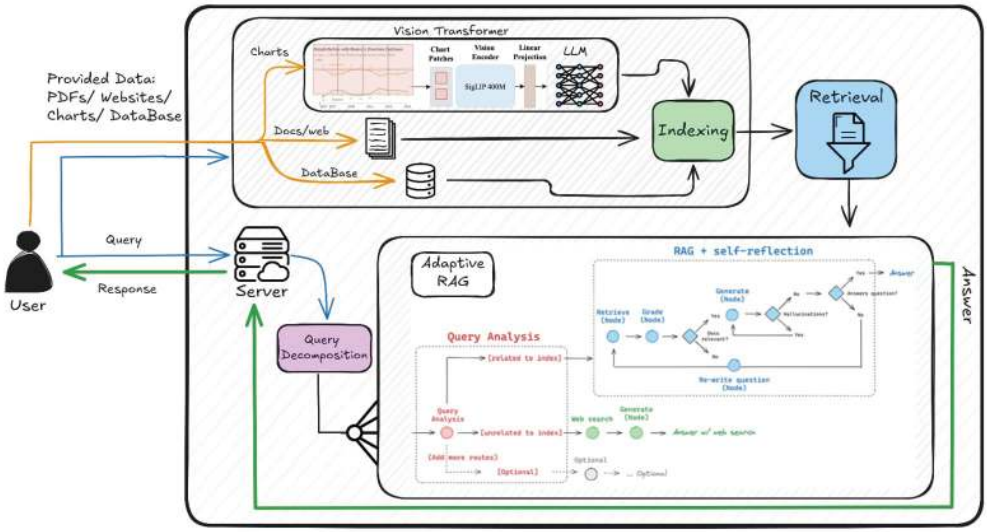

ABDULAZIZ ALOMAIR


OMAR ABDALLA


OMAR HEJAZI

In response to the growing demand for innovative technological solutions, this project introduces an advanced AI tool designed to streamline data analysis and enhance data-driven decision-making for professionals in finance, business consulting, taxation, law, and other data-intensive fields. The tool features a fine-tuned Vision-Language Model (VLM) that excels at interpreting and analyzing charts, graphs, and other visual data representations. Complementing the VLM is a Retrieval-Augmented Generation (RAG) system, which processes and synthesizes unstructured data from a variety of sources, including documents, reports, and financial statements.

By integrating these cutting-edge technologies, the AI tool empowers professionals to efficiently extract and interpret insights from both structured visual data and complex unstructured information. This seamless integration not only improves the accuracy and speed of data analysis but also significantly enhances the quality of decision-making across industries. The project aims to revolutionize how professionals interact with and leverage large datasets, providing them with a robust, AI-driven solution tailored to the demands of modern data analysis.


Vision Transformer
Charts
Chart Patches
Vision Encoder
Linear Projection
LLM
SigLIP 400M
Provided Data:
PDFs/ Websites/
Charts/ DataBase
Docs/web
DataBase
Indexing
Retrieval
Query
Response
User
Server
Adaptive RAG
RAG + self-reflection
Query Decomposition
Query Analysis
Answer

A COST-EFFICIENT, LLM ROUTING SYSTEM


HAMZA ALKAF

ABDULRAHMAN SAAD

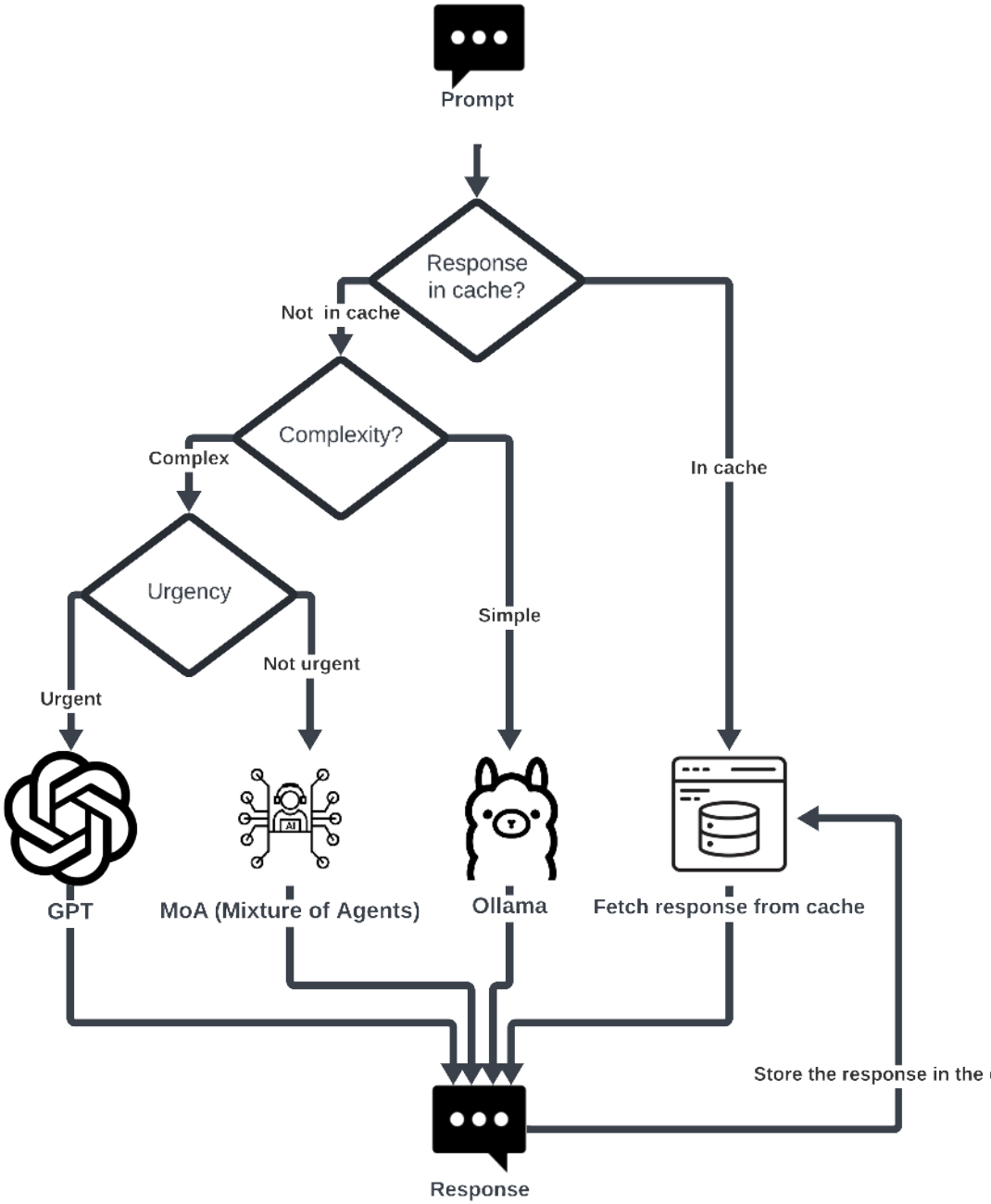

RAYAN GHABASHI


IYAS BAROUD


MOHAMMED BADAWI

ABDULQADIR BARIYAH

In this project, we offer a technique for querying LLMs while maximizing quality, speed, while maintaining minimum cost. First, a semantic cache is established to handle frequent prompts, saving time and money. Then, a router directs simple prompts to low-cost local models like llama, further maintaining the cost. Lastly, if the prompt is set to be urgent by the user, it is passed to a powerful, fast, but costly model such as GPT4o. If the request is not urgent, it is routed to a powerful free model, such as Mixture of Agents, which produces high-quality but slower results. In this study, we discuss our system's architecture as well as the results, which demonstrate the tremendous benefit of combining such basic components. Finally, the report concludes with citations to additional resources.


Prompt
Response in cache?
Not in cache
In cache
Complexity?
Complex
Simple
Urgency
Urgent
Not urgent
GPT
MoA (Mixture of Agents)
Ollama
Fetch response from cache
Store the response in the
Response

(PNEUMONIA DIAGNOSIS ASSISTANT) - CUSTOMIZED HYBRID DEEP LEARNING APPROACH FOR SYMPTOM-DRIVEN PNEUMONIA DIAGNOSIS


NOUF ALSHAWOOSH

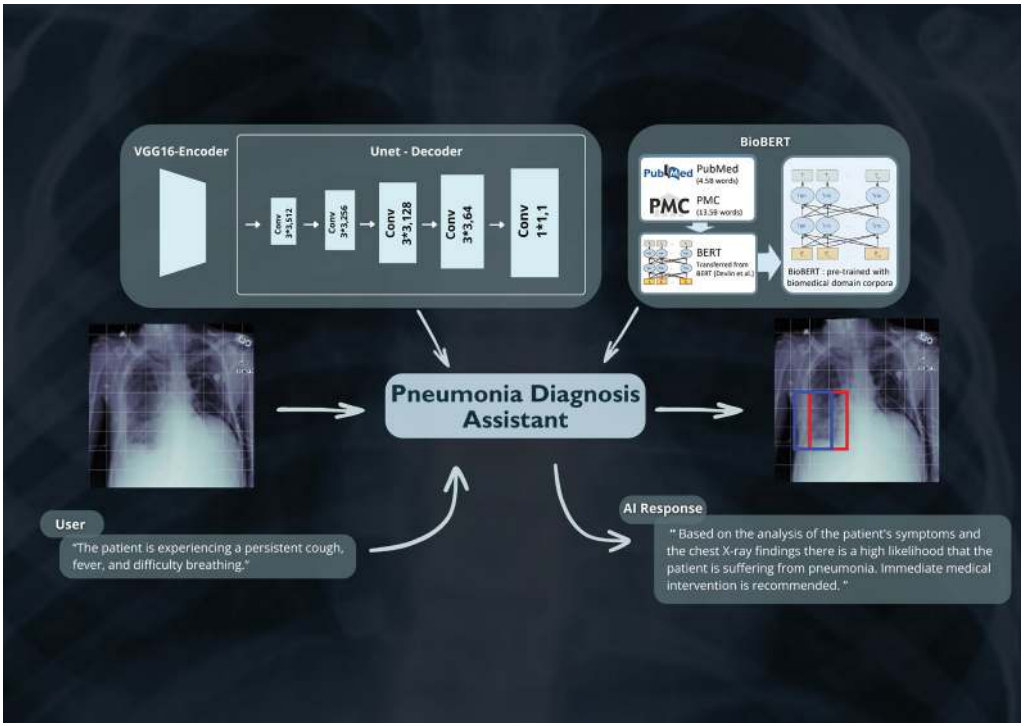
MAHA ALSAYED


RAHAF ALERWI


SAMAH HUMOOD ALHILALE

This project presents a hybrid deep-learning approach for diagnosing pneumonia, integrating image and text-based data. A VGG16-based U-Net model was developed to analyze chest X-rays, effectively detecting pneumonia with high precision. Simultaneously, a pre-trained BioBERT model was employed to extract symptoms and relevant medical history from patient records. Although BioBERT was used without fine-tuning, it served as a baseline for symptom analysis, with future work aimed at fine-tuning to enhance accuracy.

The integrated model demonstrated a strong performance, achieving an Intersection over Union (IoU) of 0.88 and an accuracy of 95%. This dual-modality approach allows for a more accurate diagnosis by combining the strengths of image segmentation and natural language processing. showcasing the potential of hybrid models in improving diagnostic tools for healthcare professionals, offering faster, more reliable, and context-aware medical diagnosis.


VGG16-Encoder
Unet - Decoder
Conv 3*3,512
Conv 3*3,256
Conv 3*3,128
Conv 3*3,64
Conv 1*1,1
BioBERT
PubMed (4.5B words)
PMC (13.5B words)
BERT
Transformed from BERT (Devlin et al.)
BioBERT : pre-trained with biomedical domain corpora
Pneumonia Diagnosis Assistant
User
"The patient is experiencing a persistent cough, fever, and difficulty breathing."
AI Response
" Based on the analysis of the patient's symptoms and the chest X-ray findings there is a high likelihood that the patient is suffering from pneumonia. Immediate medical intervention is recommended. "

RECOGNITION OF RECYCLABLE MATERIALS USING REAL-TIME OBJECT DETECTION


NAYEF YOSEF ALJHANI

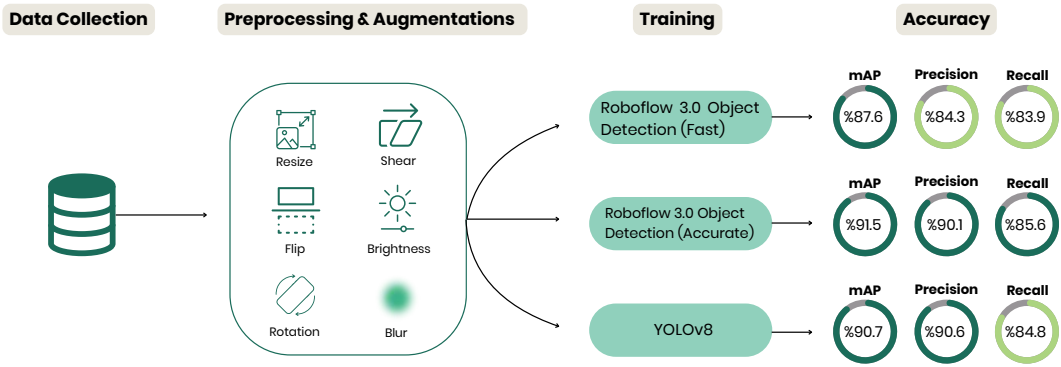

RENAD ABDULKHALEQ ALSHAMRANI


LAURA ISMAIL FATTA


NSREEN HUJJATULLAH


WALEED ALBISHRI

Recycling is a critical component of sustainable development, vital for reducing waste and conserving resources. In alignment with Saudi Arabia's Vision 2030, this project aims to address inefficiencies in current recycling methods by developing a real-time object detection and classification model for recyclable materials. Multiple models were tested, with YOLOv8 proving to have the best precision. The model automates material identification, improving the speed and accuracy of waste management processes. Future work will focus on integrating weight prediction capabilities and expanding the model's ability to classify more material types. These improvements aim to enhance recycling efficiency, reduce waste, and promote sustainability.


Data Collection
Preprocessing & Augmentations
Training
Accuracy
Resize
Shear
Flip
Brightness
Rotation
Blur
Roboflow 3.0 Object Detection (Fast)
Roboflow 3.0 Object Detection (Accurate)
YOLOv8
mAP
Precision
Recall
%87.6
%84.3
%83.9
%91.5
%90.1
%85.6
%90.7
%90.6
%84.8

47

FIREFLY AUTONOMOUS UAV SWARMS FOR WILDFIRE COMBAT UTILIZING SWARM INTELLIGENCE



ABDERRAHMANE BALAH



MOHAMMED ALWALEED ABUSHWARIB



AHMED ABDULLAH ALKHABBAZ



MUHAB MOTASEM ABUBAKER



ABDULHADI IBRAHIM ZUBAILAH

Combating wildfire outbreaks introduce a devastating financial and logistical burden on firefighting systems, with overall damage cost in 394-893B$ and 9-63K$/hr for indirect aerial costs. Our work aims to integrate the usage of multi-agent deep reinforcement learning to autonomously combat wild fires with drones. Our current work showcases the feasibility of such system on a discrete small case to be later extended into realistic 3D physics-based simulations.

FireFly Autonomous UAV Swarms for Wildfire Combat Utilizing Swarm Intelligence

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

أكاديمية كاوست
KAUST ACADEMY

Abstract

Combating wildfire outbreaks introduce a devastating financial and logistical burden on firefighting systems, with overall damage cost in 394-893B$ and 9-63K$/hr for indirect aerial costs. Our work aims to integrate the usage of multi-agent deep reinforcement learning to autonomously combat wild fires with drones. Our current work showcases the feasibility of such system on a discrete small case to be later extended into realistic 3D physics-based simulations.

Methodology

Our RL-based solution has 2 key elements:

- Simulation environment to handle the spread of the fire dynamics, and agent movement
- Neural network used to learn optimal firefighting strategies for the agents eithin the environment

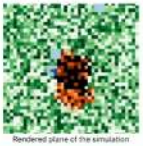
Environment

Observation Space We construct a multi-agent MDP, with assumption of full observability for each agent. a single agent observation is a 6 layered volume, with each spanning a size of $N \times N$. The layers semantics are as following:

1. Agent own position, indicated by a one in a zeroed array
2. Other agent positions
3. state S of the tree cells in the grid, where each cell can be $S \in \{healthy, burning, burnt, retardant\ healthy, retardant\ burning, retardant\ burnt\}$
4. The vegetation density, with each cell density in range [0-1]
5. The Y coordinate of the wind vector, extrapolated over the plane
6. X coordinate of the wind vector, extrapolated over the plane

Action space each agent can:

1. Move left
2. Move top-left
3. Move top
4. Move top right
5. Move right
6. Move bottom-right
7. Move bottom
8. Move bottom-left
9. Apply fire-retardant (3x3) fire retardant splash



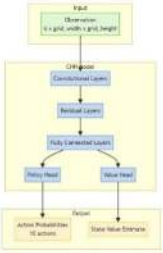
Rendered plane of the simulation

Reward Definition: At each timestep t, the agent is rewarded negatively for the number of currently burning trees. It is also, rewarded positively when applying retardant with a scaling factor that uses the euclidean distance between the agent and the nearest burning cell position. The reward for applying the retardant decays over time to incentivize the agent to put down fire as fast as possible

Architecture

Unlike all previous DQN-based solutions, we used a gradient-based solution. Namely, Multi-agent proximal policy (MAPPO), as we believe it can adapt to the rapid non-stationarity of the environment better than off-policy methods.





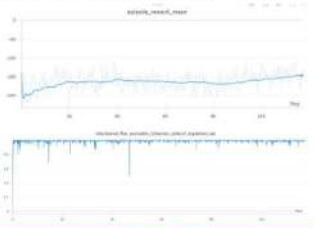
Our policy/critic networks use a 3-layered CNN backbone to process the features of the agent observation and produce action log probabilities and Q-value respectively.

Training

We used Ray's RLlib to train the agents on parallel rollout workers using 4 CPU cores, and a single cloud-hosted P100 GPU. We performed 100+ runs to test different hyperparameters and reward definitions to end with the best combination for highest episodic mean that also attempts to maximise the set objective (minimum number of burnt trees)

Results

After multiple iterations of the agent, it was able to converge on a decent reward. The critic value function was able to capture variance in the observations leading to highly explained variance, and upon inspection of agent performance recording, it was able to perform much better in encircling fire than a random agent



Future Work

We can extend our work to use partial observations with a communication channel to update the agents' history. Train the model in continuous 3D space with physics-based models like KAUST mesoscale. We can also attempt to use Model-based algorithms like DreamerV3 to allow for more robust prediction of fire spreading patterns by our agents.

References

(1) - Alexandridis, A., Vakalis, D., Siettos, C. I., & Bafas, G. V. (2008a). A cellular automata model for forest fire spread prediction: The case of the wildfire that swept through Spetses Island in 1990. Applied Mathematics and Computation, 204(2), 191–201. https://doi.org/10.1016/j.amc.2008.06.046

(2) - Hakan, R. N., & Sahraoui, H. (2018). Distributed deep reinforcement learning for fighting forest fires with a network of Aerial Robots. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). https://doi.org/10.1109/iros.2018.8593539

Team | Muhab Abubaker – Mohammed Abushwarib – Ahmed Alkhabbaz – Abdulhadi Zubailah – Abderrahmane Balah **Supervisor | Fares Fourati**

ZERO-SHOT MULTIMODAL VIDEO QUERYING WITH ADAPTIVE CANDIDATES RANKING


MOHAMED ELTAYEB


OSAMAH SARRAJ


MOHAMMED BREEMO

TAHA ALSHATIRI

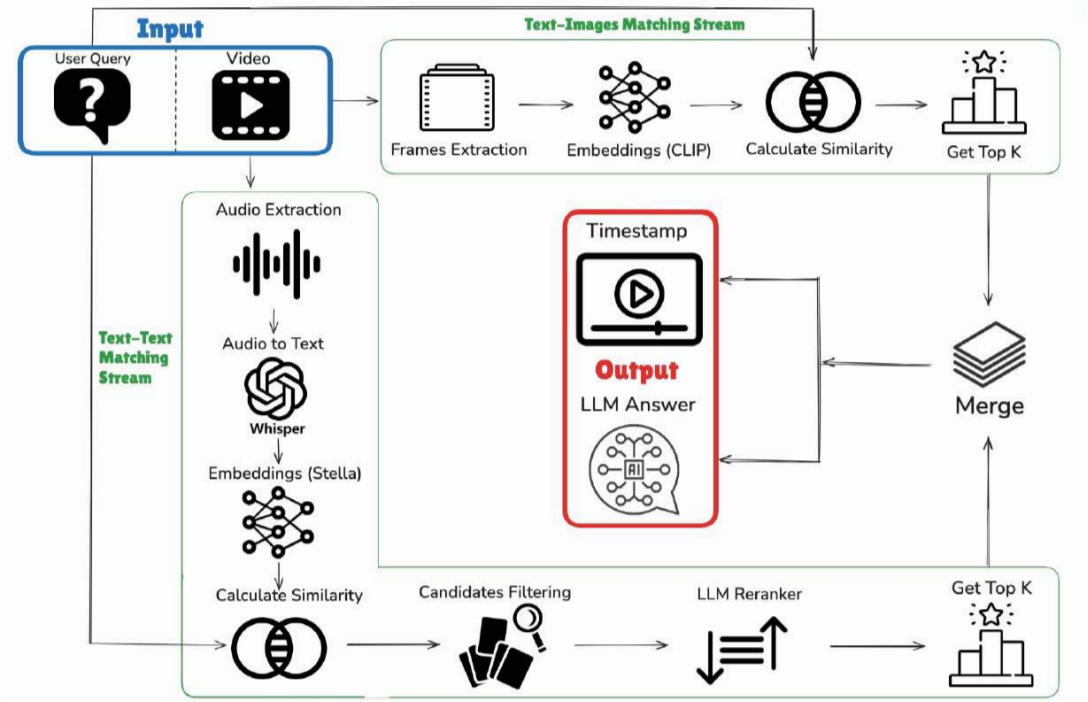

MOHAMMED ALAWI


ABDULRAHMAN FARAHIDI

Video content retrieval involves searching and extracting relevant video sections based on user queries.

We introduce a multimodal video querying system that integrates text and image data to improve accuracy and efficiency. The system uses dynamic candidates generation and ranking techniques to improve response accuracy in zero-shot scenarios, where models haven't been trained on specific datasets.

Our system has set a new state-of-the-art benchmark on YouCook2 benchmark dataset.


Input
User Query
Video
Text-Images Matching Stream
Frames Extraction
Embeddings (CLIP)
Calculate Similarity
Get Top K
Audio Extraction
Audio to Text
Whisper
Embeddings (Stella)
Calculate Similarity
Text-Text Matching Stream
Timestamp
Output
LLM Answer
Merge
Candidates Filtering
LLM Reranker
Get Top K

LEVERAGING LANGUAGE MODELS FOR ISLAMIC QUESTION & ANSWERING SYSTEM


ASMA MAHDI


LINA ALQAHTANI


REEMA ALJALAJEL

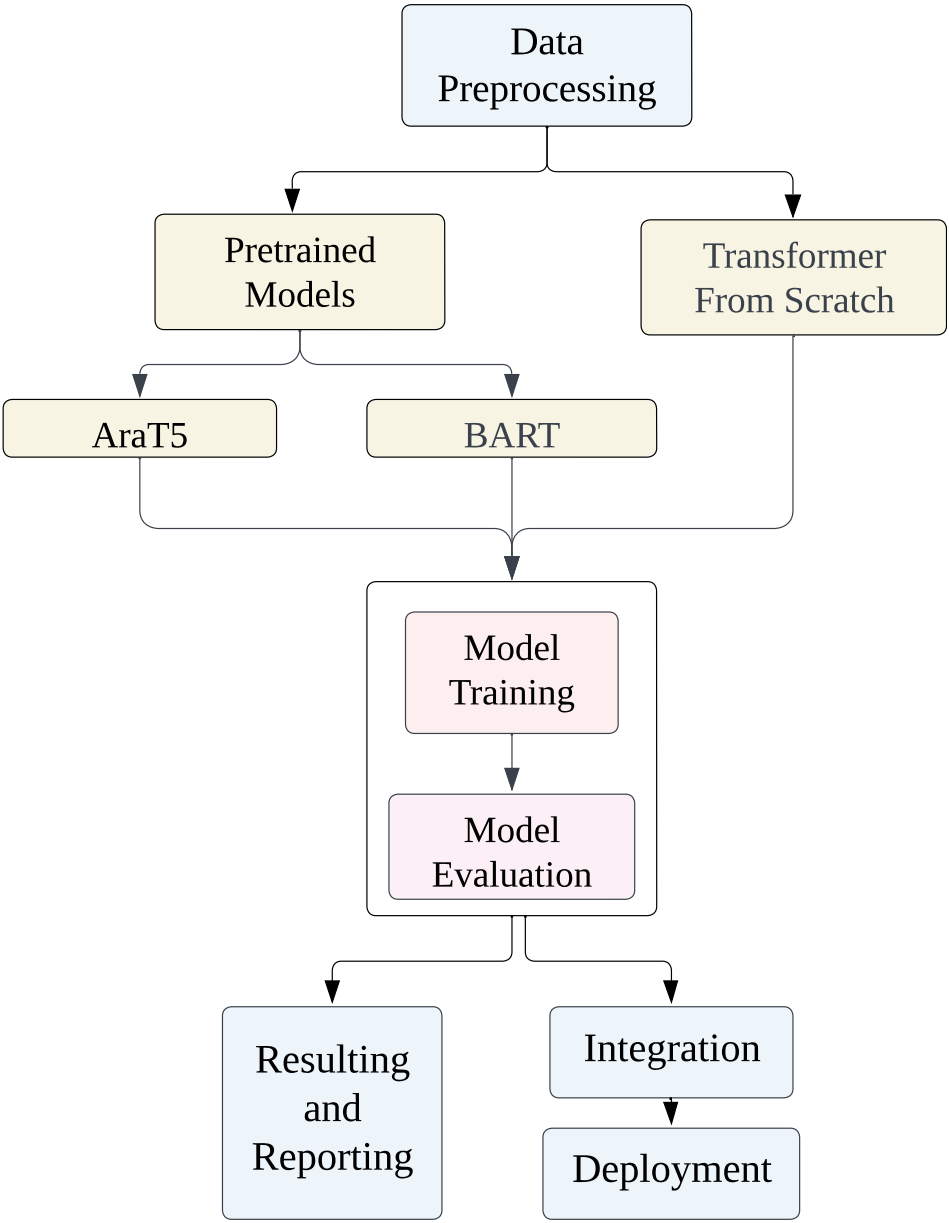

LAMA AL-QARNI


HAIFA ALTHOBAITI

Recent advancements in Natural Language Processing (NLP) have significantly enhanced the capabilities of large language models (LLMs) in understanding and generating human-like text. This progress offers promising opportunities in various domains, including the development of automated question-answering systems.

In the context of Islamic question answering, these systems face unique challenges due to the complexity and sensitivity of religious texts. This research explores the potential of leveraging state-of-the-art LLMs, specifically AraT5 and BART, fine-tuned on Islamic texts, to generate accurate and contextually relevant answers to questions about Islamic jurisprudence, theology, and ethics. A baseline transformer model that was built from scratch was also developed to provide a comparative evaluation of these advanced models.

The fine-tuning process involved adapting AraT5, a model tailored for Arabic text, and BART, known for its robust text generation capabilities, to a dataset of Islamic fatwas.


Data Preprocessing
Pretrained Models
Transformer From Scratch
AraT5
BART
Model Training
Model Evaluation
Resulting and Reporting
Integration
Deployment

MEDTWIN: BIOMEDICAL KNOWLEDGE GRAPHS FOR EARLY PROGNOSIS OF DIABETES USING DIGITAL TWIN TECHNOLOGY



LEENA ABDULLAH ALSHEHRI



SAJA SHABIB ALHIRBISH

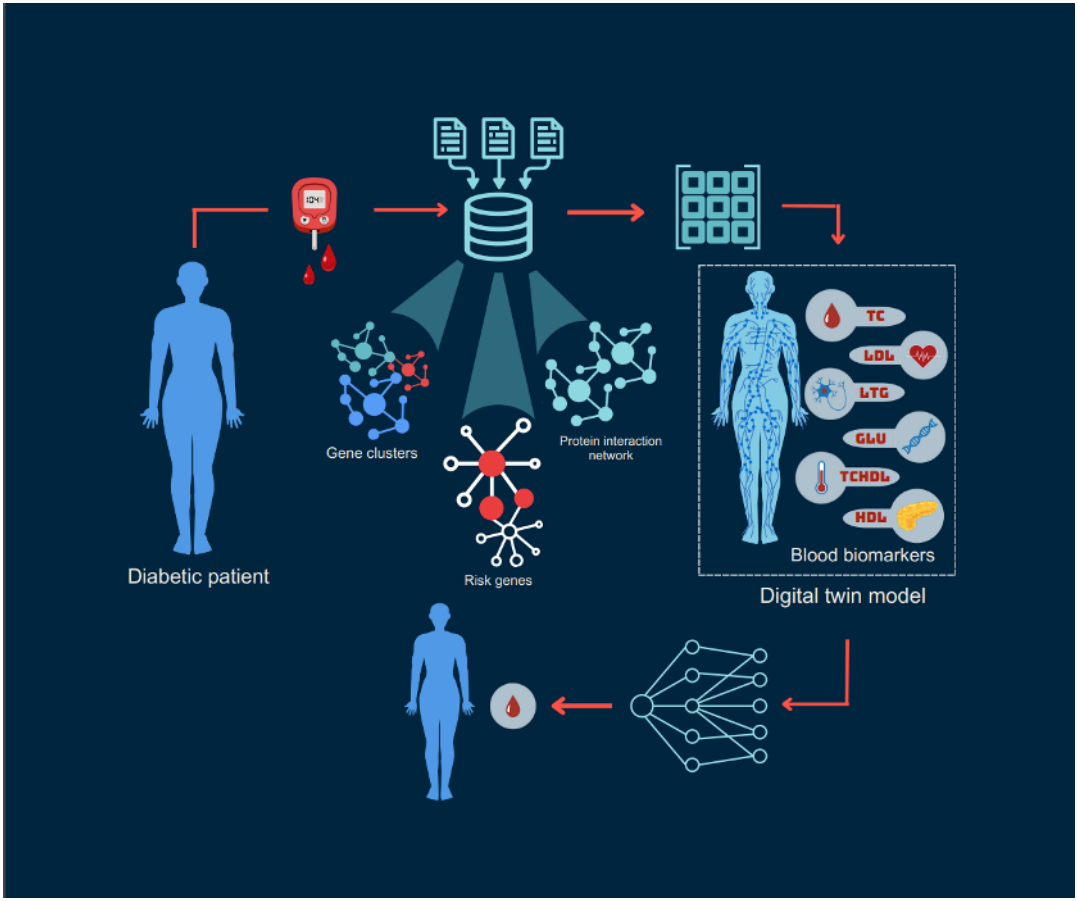


MOZN ATEF ALSHEHRI



KHADIJA CHOKRI DRIRA

The MedTwin project focuses on developing a digital twin model for the early prognosis of diabetes by integrating genomic data with blood biomarkers into a comprehensive knowledge graph. From this knowledge graph, the embeddings of biomarkers are generated, which are then used to predict the severity of diabetes in patients. Leveraging AI-driven approaches, the project enhances the precision of personalized healthcare, allowing for tailored interventions that improve patient outcomes and optimize healthcare resources. The project validates the digital twin concept, demonstrating its potential for broader application across various chronic diseases. Future work includes incorporating real-time monitoring, expanding the model to additional diseases, and integrating drug data to predict effective treatments.


Diabetic patient
Gene clusters
Protein interaction network
Risk genes
TC
LDL
LTG
GLU
TCHDL
HDL
Blood biomarkers
Digital twin model

REAL-TIME INTELLIGENT CONTROL SYSTEM TO ADDRESS FAULTS FOR DC MOTORS


MAZEN ABDULRAHIM

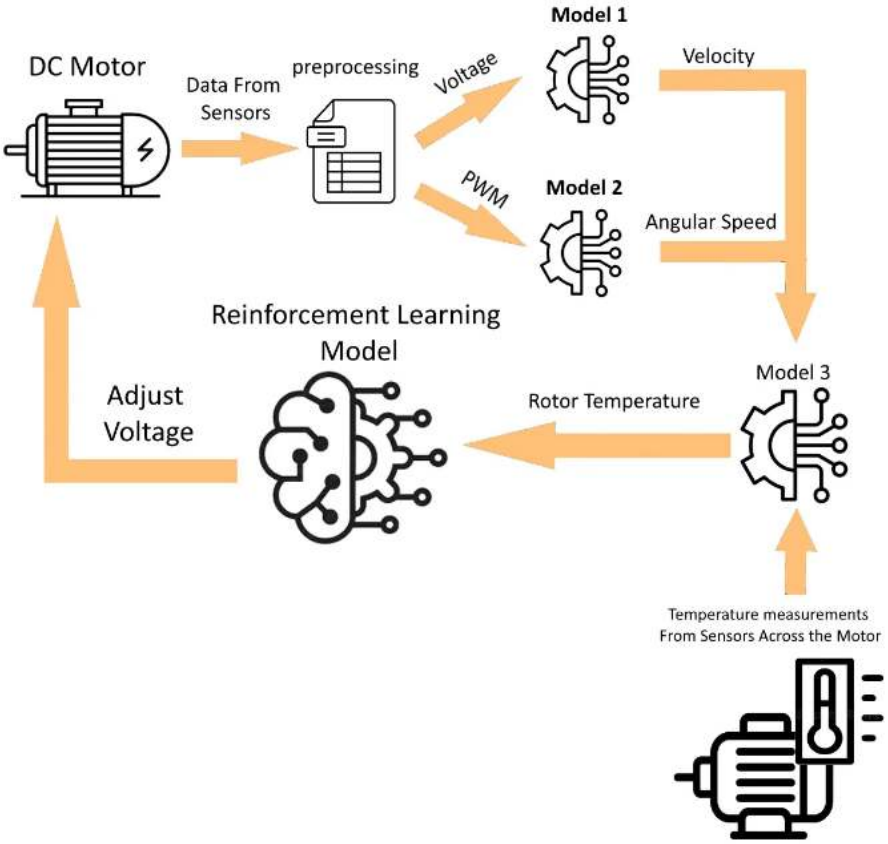

RAED ALMOZEL


KHALID MARI AL DHAFEERI

This project focuses on enhancing the modeling of Permanent Magnet DC (PMDC) motors through the application of advanced machine learning techniques. The primary objective is to develop accurate predictive models for key motor parameters, which are essential for efficient motor control and performance optimization. Initially, we developed models to predict motor velocity based on voltage inputs.

This involved training machine learning algorithms on a dataset comprising various voltage inputs and their corresponding motor velocities. Additionally, we created models to estimate rotor angular
velocity from pulse width modulation (PWM) signals. These models were trained using data that mapped PWM signals to the resulting angular velocities of the rotor.

The predictive models for motor velocity and rotor angular velocity were then integrated to estimate rotor temperature. Rotor temperature is a critical parameter that significantly affectsmotor efficiency and longevity, yet it is challenging to measure directly due to the nonlinear behavior of PMDC motors.


DC Motor
Data From Sensors
preprocessing
Voltage
PWM
Model 1
Model 2
Velocity
Angular Speed
Reinforcement Learning Model
Adjust Voltage
Rotor Temperature
Model 3
Temperature measurements From Sensors Across the Motor

HURRICANE TRAJECTORY PREDICTION


KENNETH ROLLE


JOSHUA CURRY


NIKITA JOHNSON


MATTHEW BRAYNEN


JEFF JOSEPH

Our project focuses on predicting hurricane trajectories by analyzing storm data.

Using latitude and longitude coordinates, wind speed, pressure, direction, and distance recorded at 6-hour intervals, we aim to forecast the storm's path over the next 48 hours.

The goal is to enhance forecasting accuracy and improve preparedness for hurricane impacts.

Atlantic Hurricane Paths Nestor_2019



AUTOMATED CLASSIFICATION OF PARKINSON'S GAIT SEVERITY THROUGH ACTIVITY RECOGNITION



JAWAD KHAN
NAWAB ALI



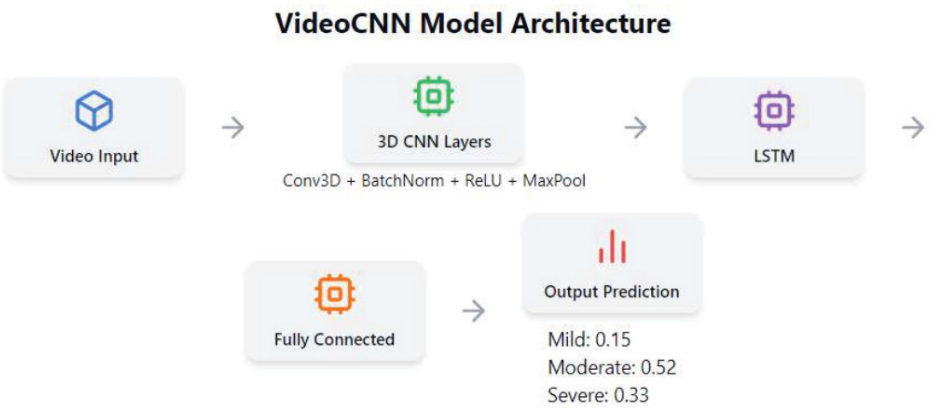
LAMA SAEED
AYASH



JANA SAAD
MATHKAR ALQAHTANI

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms, including tremors, rigidity, bradykinesia, and a particularly debilitating symptom known as freezing of gait (FoG). Accurate and early diagnosis of PD, especially the severity of gait disturbances, remains a significant clinical challenge. Traditional diagnostic methods often rely on subjective clinical assessments, which can be inconsistent and are not easily scalable. This underscores the need for objective, automated tools that can reliably assess and monitor gait disturbances in PD patients.
In this report, we propose an approach to address this diagnostic gap through the use of 3D Convolutional Neural Networks (3D-CNN) for activity recognition and classification of Parkinson's gait severity. Our method leverages 3D video analysis to automatically detect and classify gait patterns associated with different stages of PD, with a particular focus on identifying the presence and severity of FoG. The 3D-CNN architecture is trained to analyze spatiotemporal features from video data, enabling a comprehensive assessment of gait disturbances.

Our experimental results demonstrate that the proposed model achieves a classification accuracy of 89%, significantly outperforming existing methods. This high level of accuracy suggests that our approach could serve as a valuable tool for clinicians, offering a more objective and consistent means of assessing PD gait severity. Moreover, the scalability of this method presents opportunities for widespread adoption in clinical settings, potentially leading to improved patient outcomes through more precise monitoring and timely interventions.


VideoCNN Model Architecture
Video Input
3D CNN Layers
Conv3D + BatchNorm + ReLU + MaxPool
LSTM
Fully Connected
Output Prediction
Mild: 0.15
Moderate: 0.52
Severe: 0.33

AI-DRIVEN TEXT SIMPLIFICATION: UNRAVELING COMPLEX SENTENCES

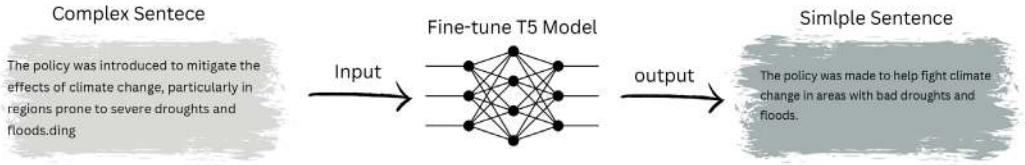


JAWAD AHMAD ALRAIMI



MUATH ABDULAZIZ

The project fine-tunes the T5 model to simplify complex scientific texts, making them more accessible without losing meaning. The model transforms technical jargon and dense academic language into clear, understandable text, aimed at broadening comprehension for researchers, students, and the general public. This work enhances accessibility to academic research and fosters interdisciplinary understanding.


Complex Sentece
The policy was introduced to mitigate the effects of climate change, particularly in regions prone to severe droughts and floods.ding
Input
Fine-tune T5 Model
output
Simlple Sentence
The policy was made to help fight climate change in areas with bad droughts and floods.

55

FARAH-CHATBOT: FINE-TUNED ARABIC RESPONSE FOR ASSESSMENT OF MENTAL WELL-BEING IN PEDIATRICS


IBRAHIM MANSOUR ALHAZMI


YAZEED SULAIMAN


MOHAMMED EMAD ALZAHRANI

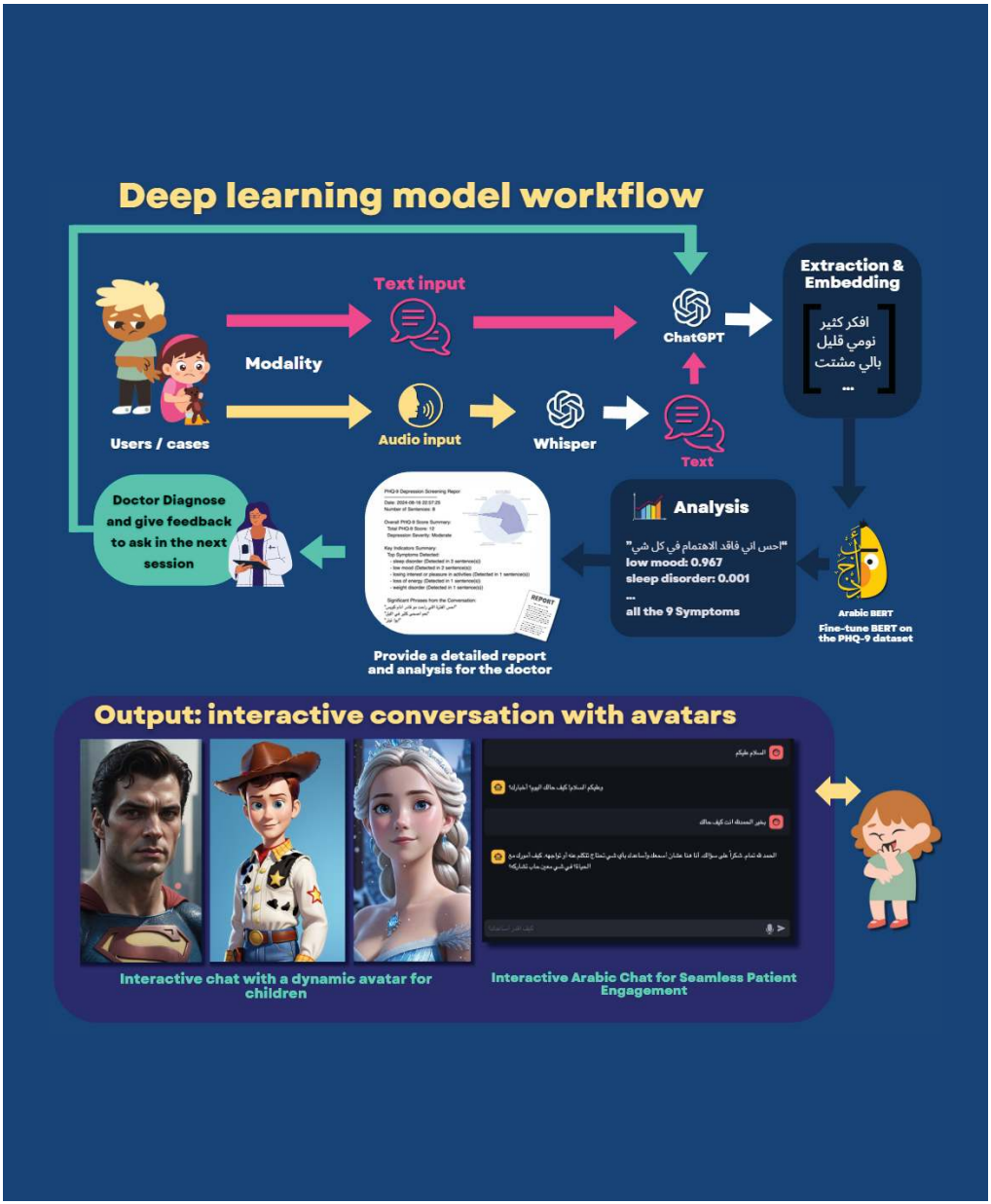

ABDULMAJEED YAHYA ALSHARAFI


MUHANNAD ABUALI

This project presents an innovative Arabic-language mental health chatbot leveraging multimodal interaction to assess and support mental well-being. The chatbot functions as a dynamic therapeutic tool, utilizing advanced natural language processing to detect and classify depression levels based on PHQ-9 criteria.

Users can interact with the chatbot through both text and voice messages, offering a seamless and adaptive user experience. Additionally, the system integrates interactive avatars, a feature tailored for children, enabling them to express emotions and thoughts to beloved characters such as superheroes.

This avatar-driven engagement creates a supportive and nurturing environment, making the mental health assessment process more approachable and emotionally resonant for younger users.


Deep learning model workflow
Text input
Modality
Users / cases
Audio input
Whisper
ChatGPT
Text
Extraction & Embedding
افكر كثير
نومي قليل
بالي مشتت
...
Doctor Diagnose and give feedback to ask in the next session
Provide a detailed report and analysis for the doctor
Analysis
"احس اني فاقد الاهتمام في كل شي"
low mood: 0.967
sleep disorder: 0.001
...
all the 9 Symptoms
Arabic BERT
Fine-tune BERT on the PHQ-9 dataset
Output: interactive conversation with avatars
Interactive chat with a dynamic avatar for children
Interactive Arabic Chat for Seamless Patient Engagement

BREACHSEEK


IBRAHIM ALSHEHRI

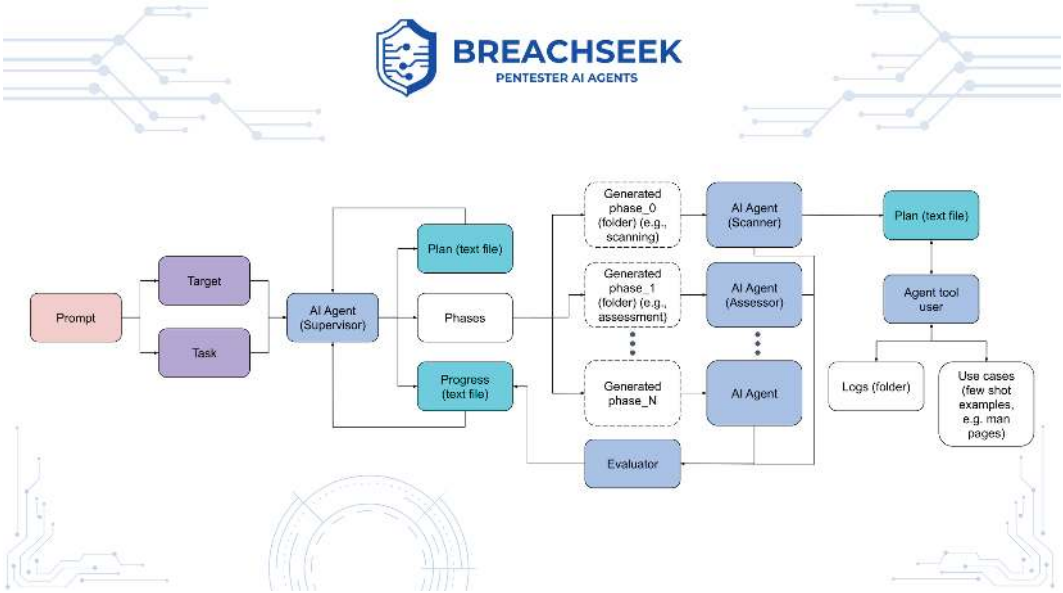

MAJED BAMARDOUF


ALAQSA AKBAR


ADNAN ALSHEHRI


ABDUL RAHMAN ALMALKI

An AI-driven multi-agent software designed to simulate cybersecurity tasks such as automated penetration testing on websites and networks. The agents utilize Large Language Models (LLMs) to identify vulnerabilities, simulate attacks, execute exploits, and generate detailed reports, leveraging knowledge from diverse data sources.


BREACHSEEK
PENTESTER AI AGENTS
Prompt
Target
Task
AI Agent (Supervisor)
Plan (text file)
Phases
Progress (text file)
Generated phase_0 (folder) (e.g., scanning)
Generated phase_1 (folder) (e.g., assessment)
Generated phase_N
Evaluator
AI Agent (Scanner)
AI Agent (Assessor)
AI Agent
Plan (text file)
Agent tool user
Logs (folder)
Use cases (few shot examples, e.g. man pages)

57

ENHANCING VIDEO FRAME INTERPOLATION: A DEPTH-AWARE UNET WITH RESIDUAL AND ATTENTION MECHANISMS


HADI AL QATTAN


DANIEL ALSADIQ


HASHEM ALMUBARAK

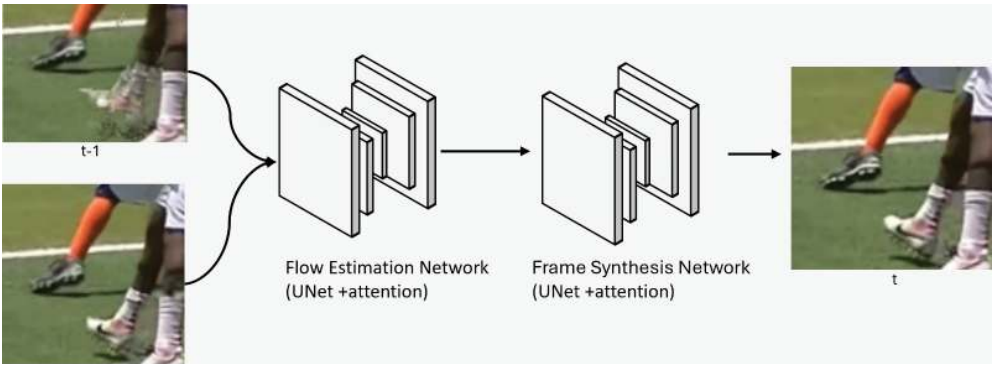
MOHAMMED AL QURAYSHAH

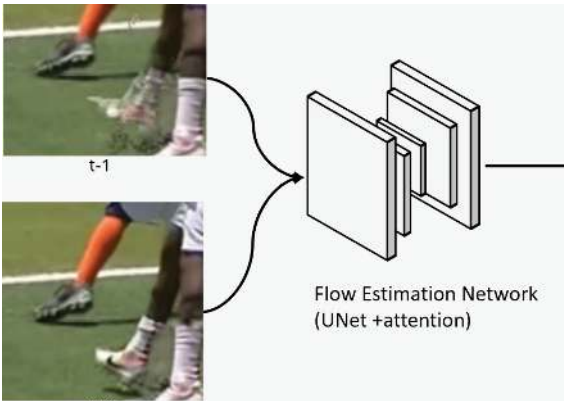

MAHDI ALABDULLATIF

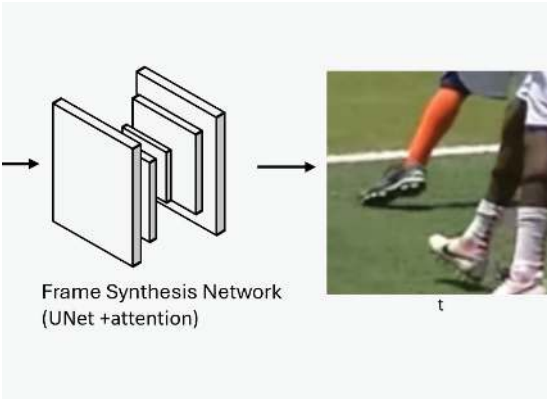
Video frame interpolation (VFI) generates intermediate frames to improve video quality and smooth motion. We introduce a Depth-Aware UNet model that enhances VFI through residual and attention blocks, improving feature extraction and focusing on critical areas of frames.

The model's encoder-decoder architecture, featuring five convolutional layers to captures higher-level features to achieve a thorough understanding of depth.

Our method focuses on leveraging depth information and highlighting key areas to generate more precise and visually coherent interpolation results.


t-1
Flow Estimation Network
(UNet +attention)
Frame Synthesis Network
(UNet +attention)
t


t-1
Flow Estimation Network
(UNet +attention)


Frame Synthesis Network
(UNet +attention)
t

58

AUGMENTING PHI-3 FOR ARABIC: GRANTING LLMS MULTILINGUAL CAPABILITY WITH MINIMAL TRAINING TIME


OMAR ABDULRAHMAN

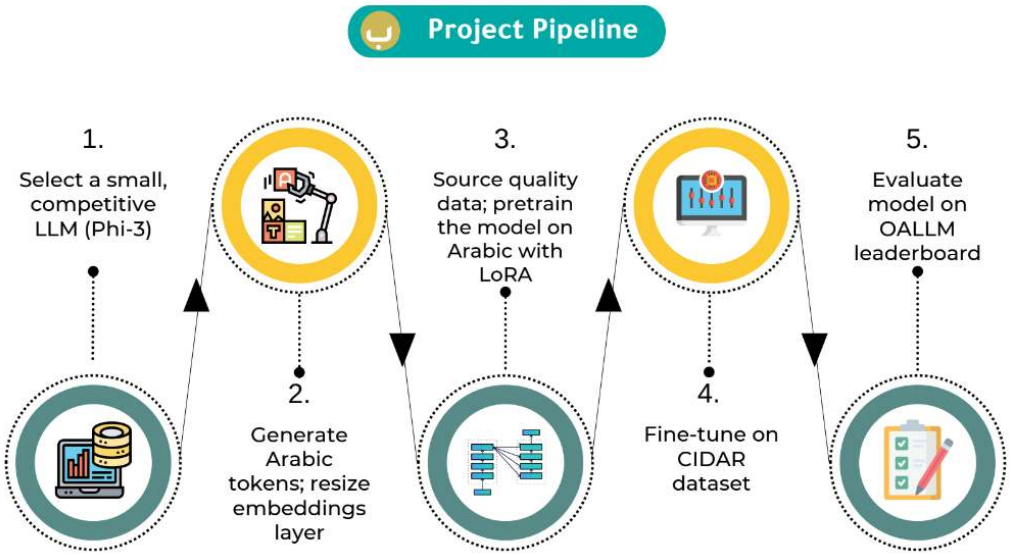

YASIR HAMDAN ALGHAMDI


AHMED ALNASSER


MOHAMMAD ATAIQ

BASSEL ESSAM KOSHAK

This project introduces a novel method for expanding the vocabulary of pre-trained Large Language Models (LLMs) for Arabic without the need for costly full pre-training. By integrating vocabulary expansion, Low-Rank Adaptation (LoRA), and targeted fine-tuning, we aim to boost model performance while minimizing computational resources. This approach sets a new benchmark for efficient language model enhancement in low-resource settings.


Project Pipeline
1.
Select a small, competitive LLM (Phi-3)
2.
Generate Arabic tokens; resize embeddings layer
3.
Source quality data; pretrain the model on Arabic with LoRA
4.
Fine-tune on CIDAR dataset
5.
Evaluate model on OALLM leaderboard

DETECTION OF BRAIN DISEASE USING ADVANCED DEEP LEARNING MODEL


BADER ALI ALSHAMRANI


BARAA MAJED ALGOMLAS

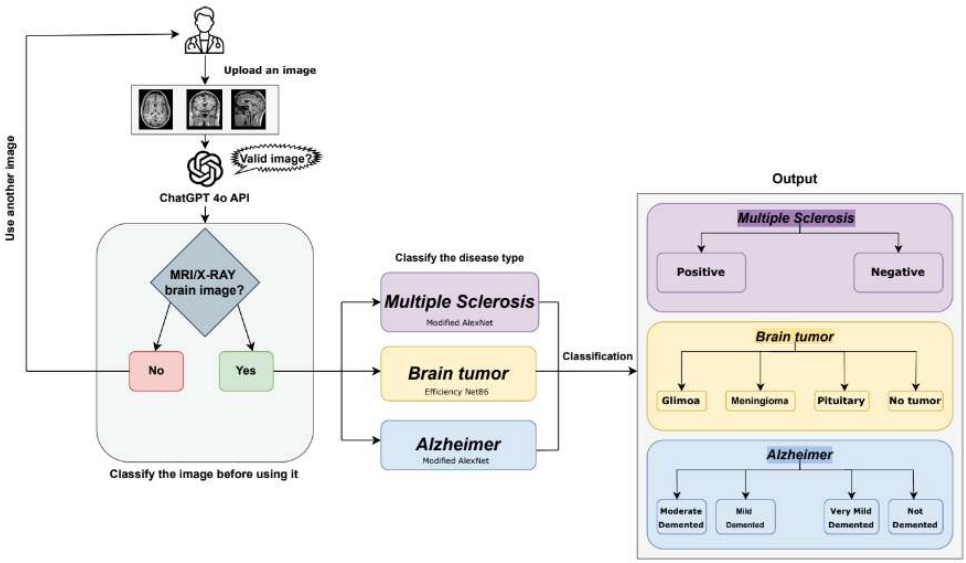

KHALID AHMED ALGHAMDI


ABDULELAH KHALAF

An AI-powered website has been developed to provide diagnostic services for brain diseases using MRI scans, focusing on Tumor Detection, Alzheimer's Disease Diagnosis, and Multiple Sclerosis Detection. The platform enables users to upload brain images for immediate analysis.

The Tumor Detection module employs the Efficiency Net86 model to classify MRI images into glioma, meningioma, pituitary tumors, or no tumor.
Alzheimer's Disease module, utilizing a modified AlexNet, categorizes images into five stages, from Non-Demented to Severe Demented, with an estimation of disease progression.
The Multiple Sclerosis module utilizes a modified AlexNet to identify brain lesions and determine whether the scan is positive or negative for multiple sclerosis.

The integration of advanced models and a robust dataset ensures precise, reliable diagnostics, aiding in the early detection and management of brain diseases.


Use another image
Upload an image
Valid image?
ChatGPT 4o API
MRI/X-RAY brain image?
No
Yes
Classify the image before using it
Classify the disease type
Multiple Sclerosis
Modified AlexNet
Brain tumor
Efficiency Net86
Alzheimer
Modified AlexNet
Classification
Output
Multiple Sclerosis
Positive
Negative
Brain tumor
Glimoa
Meningioma
Pituitary
No tumor
Alzheimer
Moderate Demented
Mild Demented
Very Mild Demented
Not Demented

THE MEETING WIZARD - UNLOCK THE POWER OF IN-PERSON MEETING USING LLM + LAM CAPABILITIES

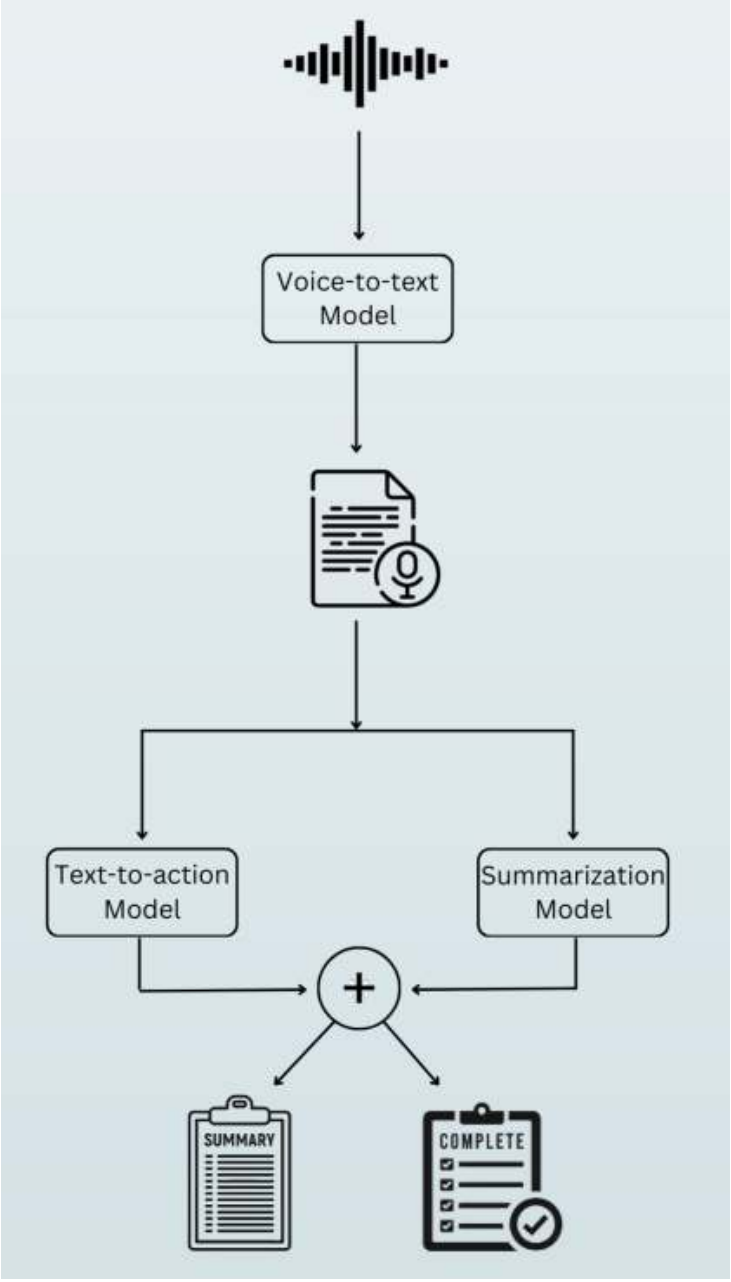

ABDULLAH ALHABIB


ABDULRAHMAN ALGHAMDI


ABDULLAH ALWEHAIBI


FAHAD ALOTAIBI

This project introduces an AI-driven system that automates the transcription, summarization, and task extraction from meetings, significantly improving business efficiency. The system streamlines task identification and management, automatically executing simple tasks through integration with workspace management tools and Google's services, allowing teams to focus on more complex actions. By enhancing workflow organization, this technology holds promise for future applications, offering the potential to further optimize operations, increase productivity, and adapt to evolving business needs.


Voice-to-text Model
Text-to-action Model
Summarization Model
SUMMARY
COMPLETE

61

FAULT DETECTION FOR INDUSTRIAL ROBOTIC ARM


ABDULHADI BASHATHI

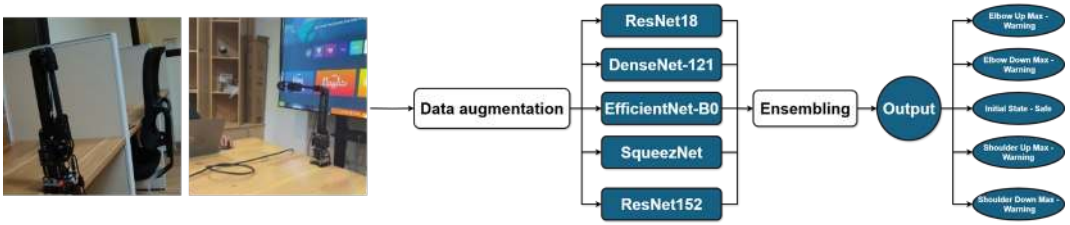

HAMED FALLATAH


ABDULAZIZ GHURAB


MOHAMMED ALSAIARI

YAQOUB WADDAD

In industrial settings, ensuring the safety and efficiency of robotic arms is crucial for minimizing risks and downtime. This research explores the use of AI to diagnose and predict faulty arm poses, offering a proactive solution to prevent hazards and enhance operational efficiency. Traditional sensor-based systems can be compromised by environmental factors, leading to undetected faults and costly inefficiencies. By applying advanced AI models for pose estimation and fault detection, this approach aims to improve the safety and reliability of robotic operations. The methodology emphasizes two critical aspects of the dataset: background consistency and object similarity to the robotic arm, ensuring a robust and accurate model for industrial applications.


Data augmentation
ResNet18
DenseNet-121
EfficientNet-B0
SqueezNet
ResNet152
Ensembling
Output
Elbow Up Max - Warning
Elbow Down Max - Warning
Initial State - Safe
Shoulder Up Max - Warning
Shoulder Down Max - Warning

LIGHTWEIGHT MODEL FOR SALIENCY PREDICTION


ABDALLAH OSAMA HAMMAD

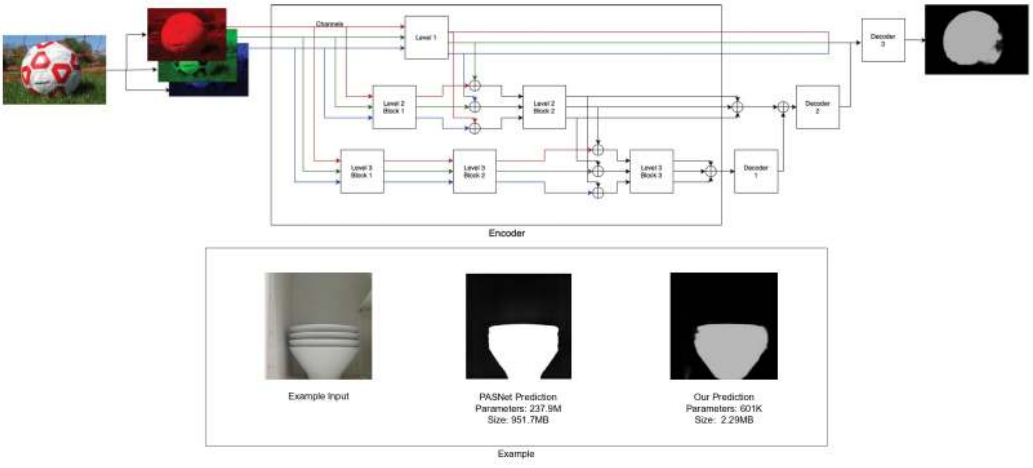

ABDULLAH AHMED BIN-TALIB


FAWAZ IBRAHIM ALBAHUSSEIN


ABDULLAH EID AL-QAHTANI


MOHAMMAD AHMED ALMATRAFI

Salient object detection (SOD) focuses on identifying the most visually prominent parts of an image, typically those that stand out due to distinct colors. In our project, we present a straightforward SOD method that leverages color separation to enhance feature extraction. This technique reduces computational demands while maintaining high performance, making it ideal for resource-constrained devices. Despite the simplicity and lightweight nature of our model, we achieved results that surpass expectations.


Channels
Level 1
Level 2 Block 1
Level 2 Block 2
Level 3 Block 1
Level 3 Block 2
Level 3 Block 3
Decoder 3
Decoder 2
Decoder 1
Encoder
Example Input
PASNet Prediction
Parameters: 237.9M
Size: 951.7MB
Our Prediction
Parameters: 601K
Size: 2.29MB
Example

BENEATH THE WAVES: LEVERAGING SAM FOR CORAL REEF DATASET GENERATION & ADVANCED VISION MODELS

SALEHAH ALJOUFI



NORAH WALEED

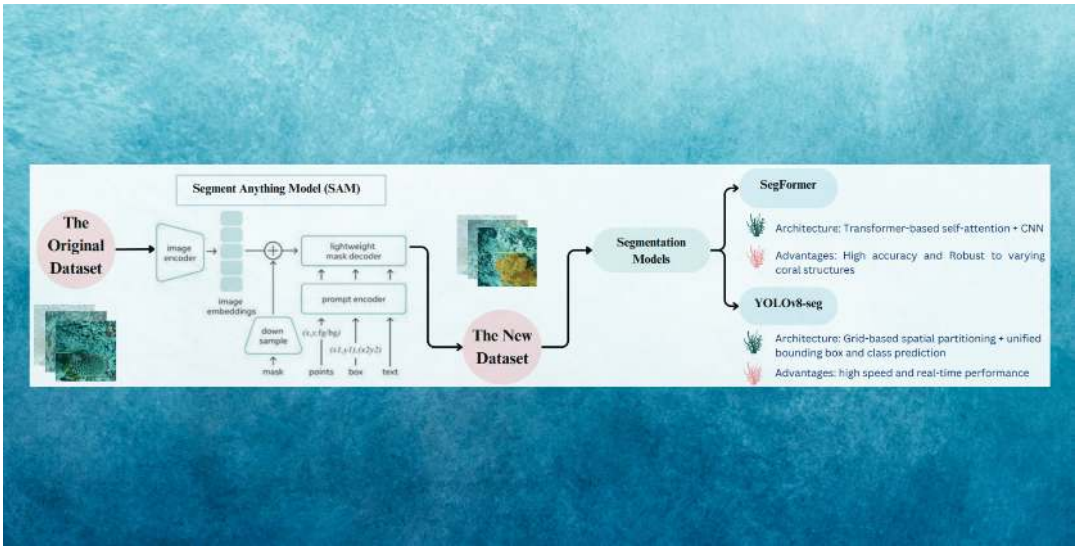
ZAHRAH ALHARZ

Inspired by the KAUST-Neom coral restoration project in the Red Sea, our work leverages advanced computer vision and deep learning to protect one of the most vital ecosystems on Earth.

Coral reefs, essential to marine biodiversity but under threat from climate change and human activity, require innovative monitoring solutions.

We developed an automated tool that uses precise segmentation to efficiently monitor coral health and support large-scale restoration.

This technology aims to enhance global conservation efforts, providing real-time insights to help safeguard the future of our oceans.


Segment Anything Model (SAM)
The Original Dataset
image encoder
image embeddings
lightweight mask decoder
prompt encoder
down sample
mask
points
box
text
The New Dataset
Segmentation Models
SegFormer
Architecture: Transformer-based self-attention + CNN
Advantages: High accuracy and Robust to varying coral structures
YOLOv8-seg
Architecture: Grid-based spatial partitioning + unified bounding box and class prediction
Advantages: high speed and real-time performance

INSIGHT: AI STOCK PREDICTION



NAIF ALZMAMI



MOHAMMED ALHARBI

NAIF ALQHATANI

This report explores the application of Deep Reinforcement Learning (RL), specifically the Deep Deterministic Policy Gradient (DDPG) algorithm, in predicting and optimizing trading decisions in the Saudi stock market. The goal is to assess the model's ability to outperform traditional strategies like buy-and-hold. The experiments reveal varied outcomes: in the best scenarios, the model mimicked the buy-and-hold strategy, achieving similar returns; in good scenarios, it managed to make profitable trades, although still close to buy-and-hold performance; in bad scenarios, it underperformed, highlighting the model's volatility; and in some cases, it failed to learn any effective trading strategy, resulting in poor performance. These results indicate that while DDPG shows promise, its current form is unstable and inconsistent, requiring further research to enhance its reliability and effectiveness in real-world trading environments.

Abstract

In recent years, Artificial Intelligence (AI) has revolutionized various industries, including finance. One of the most promising branches of AI is Reinforcement Learning (RL), which focuses on training models to make sequences of decisions. This project aims to leverage RL to predict stock prices in the Saudi stock market, enabling automated trading with optimized buy and sell decisions.

Stock trading is inherently risky and unpredictable. Traditional methods of stock prediction, such as technical and fundamental analysis, rely heavily on historical data and human judgment, which can be subjective and prone to error. This project seeks to develop an RL model that can analyze large datasets and learn from market patterns to make accurate predictions about the optimal times to buy and sell stocks.

AF-100
TACTICAL TARGET DIFFERENTIATION

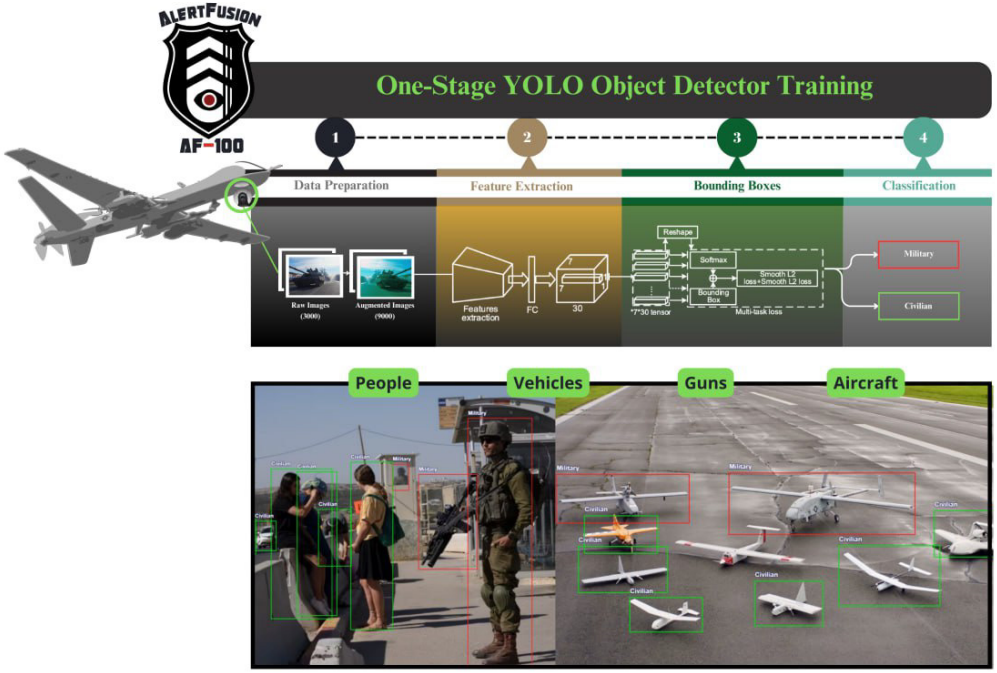
FAISAL AHMAD ALZAHRANI



AHMED IBRAHIM ALJUAID

The AF-100 Tactical Target Differentiation project focuses on developing an advanced computer vision system to accurately distinguish between military and civilian objects in real-time during military operations. Utilizing deep learning techniques like Convolutional Neural Networks (CNNs) and Transfer Learning, the system enhances target identification and classification. Object detection algorithms, including YOLO and Faster R-CNN, are employed to ensure rapid and accurate differentiation in various environments.

Image segmentation methods, like U-Net, improve detection accuracy, particularly in complex scenes. The project also explores synthetic data generation and domain adaptation to ensure robustness. The system is designed for integration into existing military reconnaissance and


ALERTFUSION
AF-100
One-Stage YOLO Object Detector Training
1
2
3
4
Data Preparation
Feature Extraction
Bounding Boxes
Classification
Raw Images
(3000)
Augmented Images
(9000)
Features
extraction
FC
30
Reshape
Softmax
Smooth L2
loss+Smooth L2 loss
Bounding
Box
*7*30 tensor
Multi-task loss
Military
Civilian
People
Vehicles
Guns
Aircraft

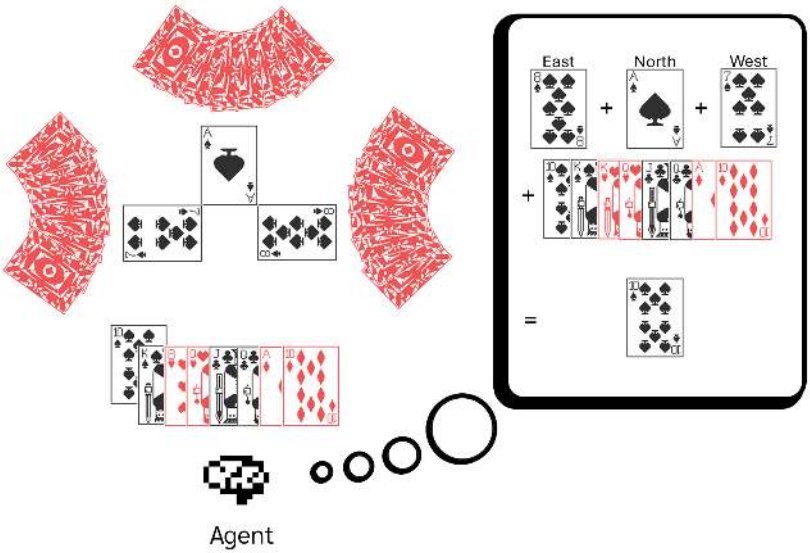
SUNAGENT


SAEED BAWAZEER


ALYAHYA AKKAM


ABDULMALIK ALMADHI

SunAgent is a project aimed at developing a specialized Reinforcement Learning (RL) agent to learn and master the Sun phase of the card game Baloot. The project involves designing and implementing the agent within a custom environment created using PettingZoo framework. To demonstrate the agent's learning and decision-making capabilities against other agents and players, we built a simple Sun game using pygame.


East
North
West
Agent

ACCELERATED POST-DISASTER RECOVERY: AN AI-POWERED SYSTEM FOR DAMAGE ASSESSMENT AND RECONSTRUCTION PRIORITIZATION

LAYAN ZAAFARANI

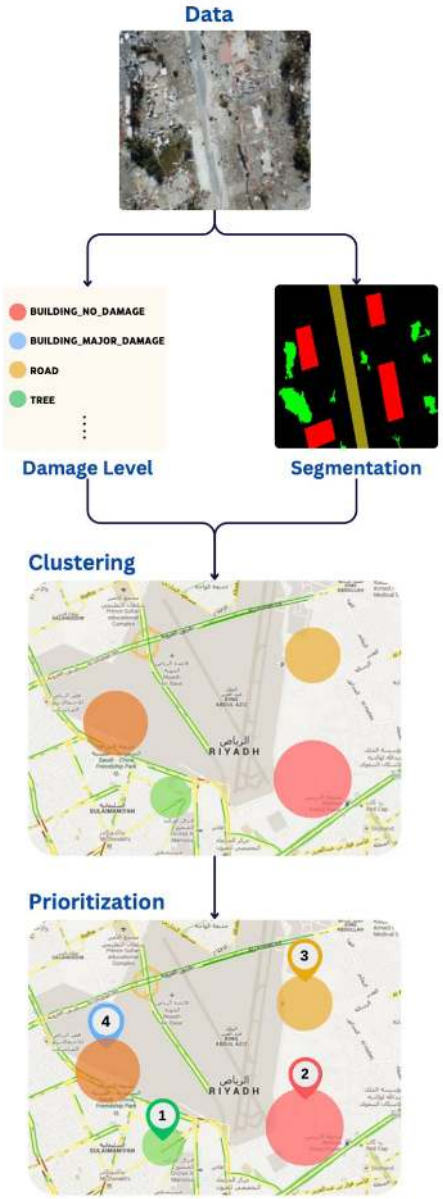
GHADI BABOUR

IMTENAN ALGHAMDI

DANAH ALMUHAYSIN

This project introduces an innovative AI-driven framework to revolutionize post-disaster recovery efforts. Leveraging UAV imagery and advanced AI techniques, our system rapidly assesses damage to buildings and infrastructure using an Attention U-Net model for image segmentation. Unsupervised learning then clusters affected areas based on damage severity, building importance, and distance, utilizing both segmentation results and Google Maps metadata.

A reinforcement learning model, employing PPO and DQN algorithms, optimizes resource allocation and prioritizes reconstruction efforts, aiming to minimize total recovery time. This integrated approach promises faster, more accurate damage assessment and data-driven decision-making, significantly reducing the time needed to restore affected communities and setting a new standard for efficient disaster response strategies.


Data
BUILDING_NO_DAMAGE
BUILDING_MAJOR_DAMAGE
ROAD
TREE
Damage Level
Segmentation
Clustering
RIYADH
Prioritization
RIYADH
1
2
3
4