**Accurate Name Extraction from News Video Graphics (284 words)**

As news consumption increasingly shifts to video, extracting key details efficiently is more crucial than ever. Identifying names embedded in graphics, such as lower-thirds and captions, presents a significant challenge due to their often brief appearance. The manual recording of these names is laborious, susceptible to errors, and time-consuming.

The Accurate Name Extraction Pipeline (ANEP) automates the name extraction process. The use of artificial intelligence facilitates the extraction of names from news videos with notable accuracy and efficiency, thereby improving information retrieval for journalists, researchers, and media analysts.

ANEP incorporates three fundamental technologies: YOLO (You Only Look Once), a deep-learning model for detecting graphical elements within video frames; Optical Character Recognition (OCR) for text extraction; and Named Entity Recognition (NER) to identify only pertinent names. ANEP's performance is assessed by comparing it to two Generative AI models, focusing on accuracy, processing speed, and effectiveness across various news formats and conditions.

The News Graphics Dataset (NGD) was developed for this thesis to detect graphical elements in news videos that contain text. It includes video frames sourced from both local and international broadcasters such as TVM, One, Net, BBC, Sky and ABC News, as well as social media-based platforms like Daily Mail, Fox News, Side Street, and Lovin Malta. Challenges including text variations, motion blur, and inconsistent fonts were mitigated through dataset refinement and optimisation of OCR.

This research emphasises the role of AI in media analysis through the development of an automated pipeline that enhances accessibility and efficiency in news reporting, thereby reducing dependence on manual transcription. The findings provide a thorough assessment of AI-driven text extraction, highlighting the relative advantages and disadvantages of traditional and generative models in the analysis of real-world news graphics.

**Accurate Name Extraction from News Video Graphics *(456 words)***

As news consumption increasingly shifts to video, efficiently extracting key details is more important than ever. Identifying names embedded in graphics, such as lower-thirds and captions, presents a significant challenge due to their brief appearance. These graphical elements vary in design, colour, and placement, making it difficult for automated systems to detect and extract relevant names. Manually recording these names is time-consuming, prone to errors, and inefficient, especially for journalists and media analysts handling large volumes of video content. Without an automated solution, critical names may be missed, affecting the accuracy of news reporting and media analysis.

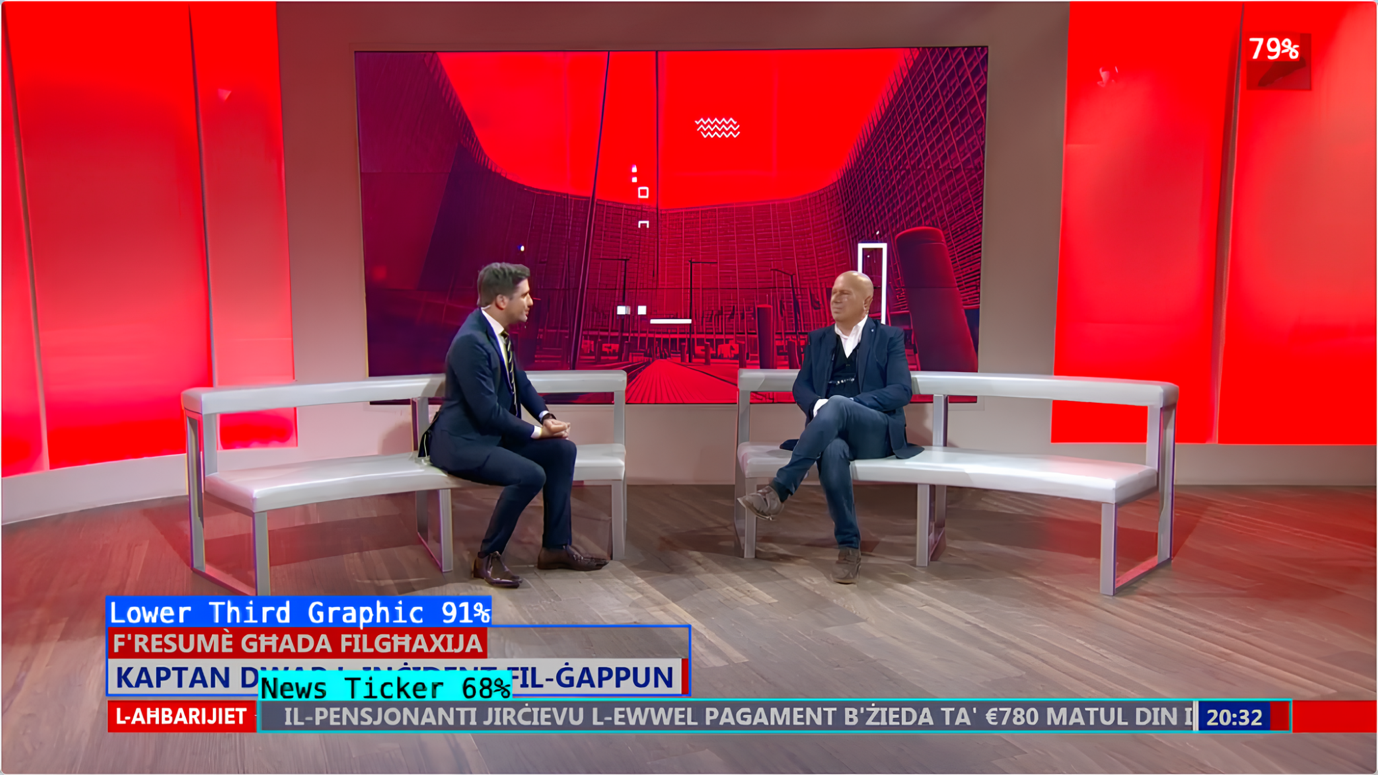
The Accurate Name Extraction Pipeline (ANEP) was developed to address this challenge. It combines three core technologies: YOLO (You Only Look Once), a deep-learning model that detects graphical elements within video frames; Optical Character Recognition (OCR), which extracts text from these graphics; and Named Entity Recognition (NER), which identifies and validates relevant names. YOLO helps isolate text-containing regions, OCR converts the detected text into a machine-readable format, and NER ensures only meaningful names are retained. ANEP’s performance is evaluated against two Generative AI models, focusing on accuracy, processing speed, and effectiveness across different news formats. These comparisons provide insight into the capabilities and limitations of different AI-driven approaches.

A crucial component of this dissertation is the News Graphics Dataset (NGD), which consists of video frames from both local and international news sources. Local outlets include TVM, One, and Net News, while international sources comprise BBC, Sky, and ABC News. Recognising the increasing role of social media in news consumption, the NGD also includes social media-based sources such as Daily Mail, Fox News, Side Street, and Lovin’ Malta. The dataset spans various video qualities, graphical styles, and text overlays, ensuring ANEP can generalise effectively.

During development, several technical challenges arose, particularly with motion blur, inconsistent font styles, and overlapping graphical elements. These issues were mitigated through dataset refinement and augmentation, pre-processing techniques, and OCR optimisation to enhance text clarity. The comparative analysis highlighted ANEP’s strengths in structured name detection, demonstrating high accuracy in controlled conditions, while Generative AI models showed flexibility in interpreting partially obscured text. The differences between these models highlight the trade-offs between structured AI approaches and more flexible, context-driven Generative AI solutions.

This research underscores the role of AI in media analysis by providing an automated pipeline that enhances the efficiency and accessibility of news reporting. By automating name extraction, ANEP supports media professionals in retrieving critical information quickly and accurately. The findings offer a detailed evaluation of AI-driven text extraction, illustrating the strengths and limitations of traditional and Generative AI models. Future work will focus on enhancing ANEP’s real-time processing capabilities, improving adaptability to evolving graphical design trends, and integrating user feedback to refine detection accuracy.

**Images & Captions *(Image1 is preferred)***



Caption: Test frame showcasing YOLOv12 Object Detection (Fast) after 77 epochs, accurately identifying different types of news graphics, including the lower third (91%) and news ticker (68%). The model demonstrates high precision in detecting key on-screen elements essential for automated media analysis.

A diagram of anep pipeline overview

AI-generated content may be incorrect.

Caption: Overview of the Accurate Name Extraction Pipeline (ANEP), detailing the sequential stages from video upload to name comparison. The system integrates AI-driven analysis with multiple extraction approaches, enabling robust identification and evaluation of names from news video graphics.