Protective Vision

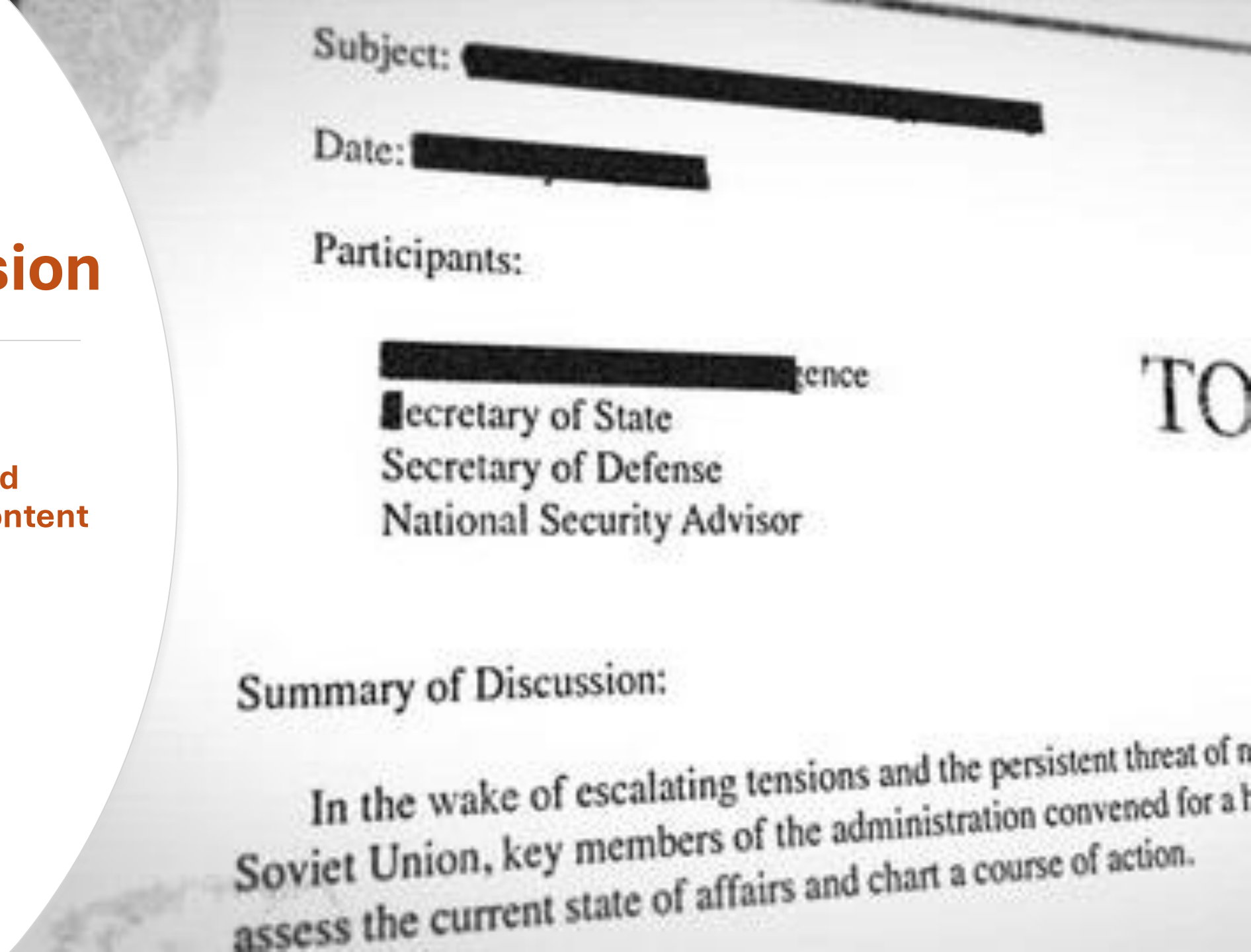
A Computer Vision Approach to Sensitive Information Detection and Hiding in Social Media Content

Presented By:

Muqeet Ahmad
Ahmad Shafique

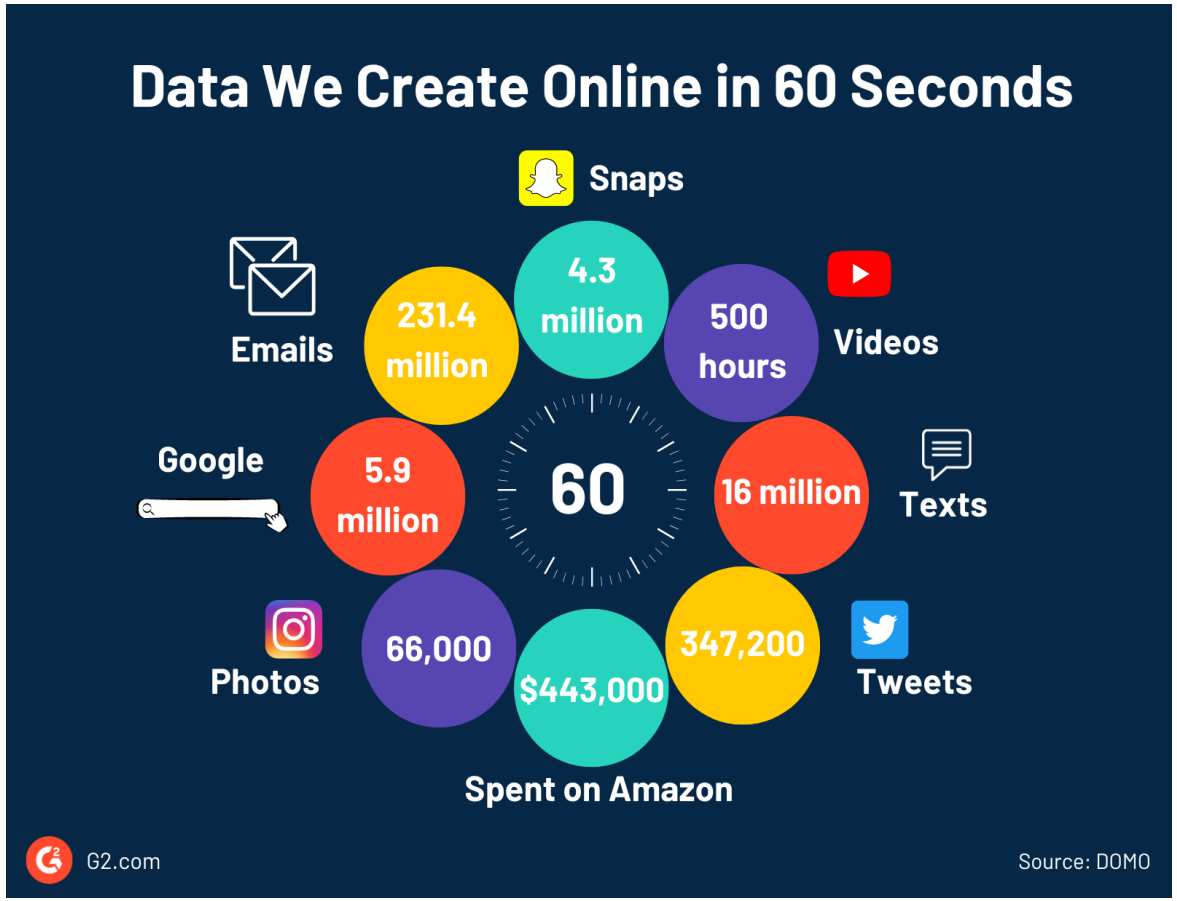
Supervised By:

Dr. Muhammad Safyan



Background

- The amount of data produced and shared online is increasing at an exponential rate, leading to growing privacy vulnerabilities.
- Sharing sensitive information on social media can result in **social engineering attacks** and **privacy breaches.**
- Your **vehicle number plate** can be misused by unauthorized individuals, potentially leading to **legal or criminal implications.**
- Your **bank card** left visible in a casual photo may expose confidential details.
- Your **contact number or email address** can be leaked unknowingly through platforms like **WhatsApp statuses or stories.**



Literature Review

- Current systems focus only on **detection**, not **redaction**
- No single tool handles **multiple types** of sensitive info
- Most systems are cloud/desktop-based — not **mobile-friendly**
- Detection without masking risks **exposure**
- **One-click solution** to redact everything

Dataset, Augmentation and Training

Type	Original	Augmented	Total	Training	Testing	Validation
Vehicle Plate	6200	5589	11789	8252	1768	1769
Banking Card	1350	1215	2565	1795	385	385

Original Image

After Augmentation

Model Training & Results

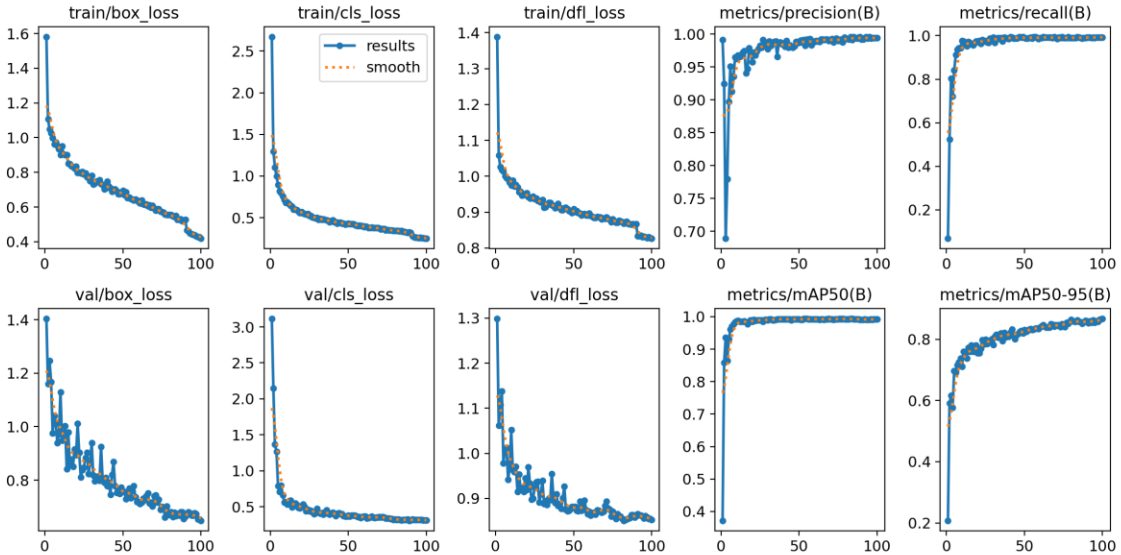
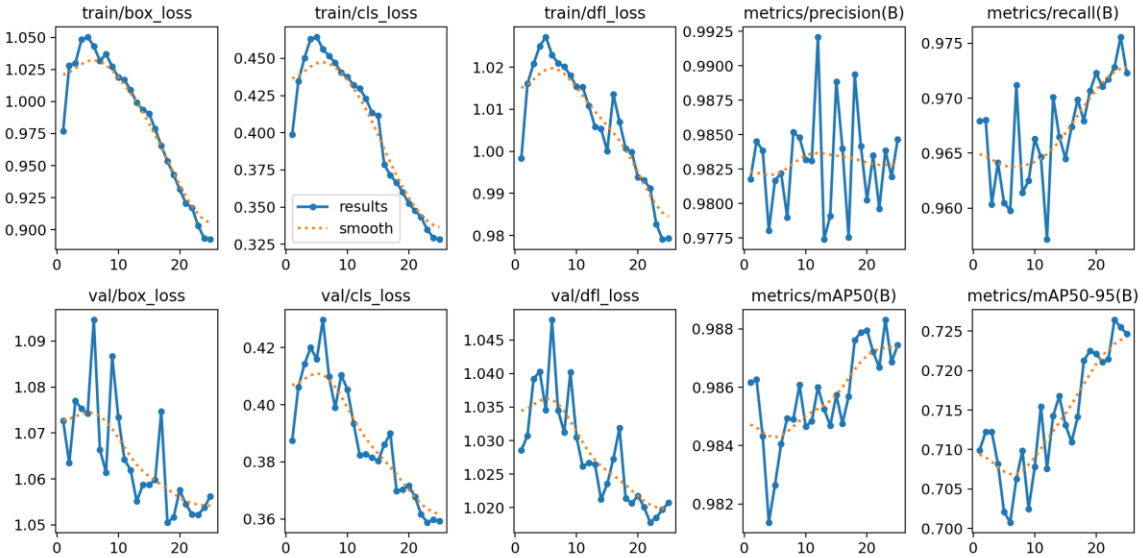
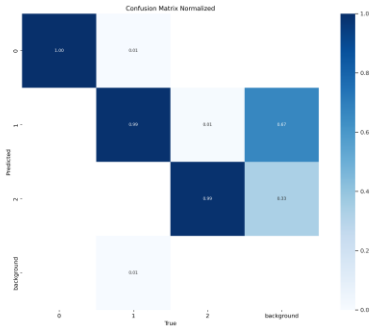
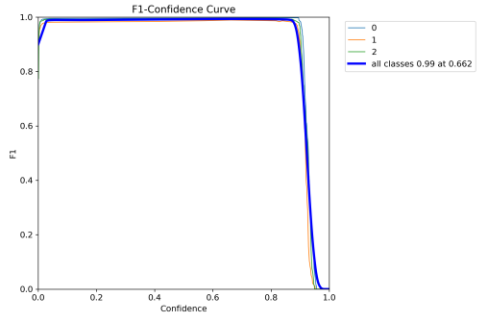


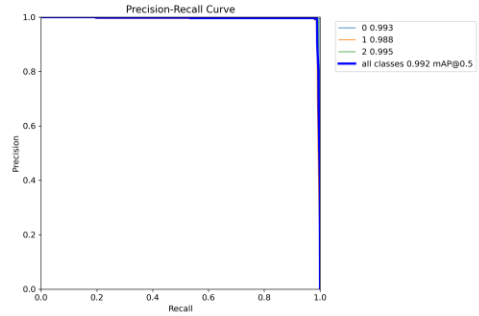
Plate Model

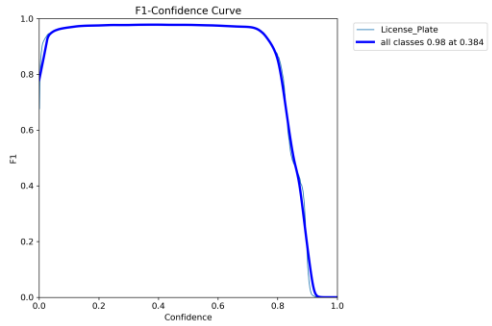
Card Model

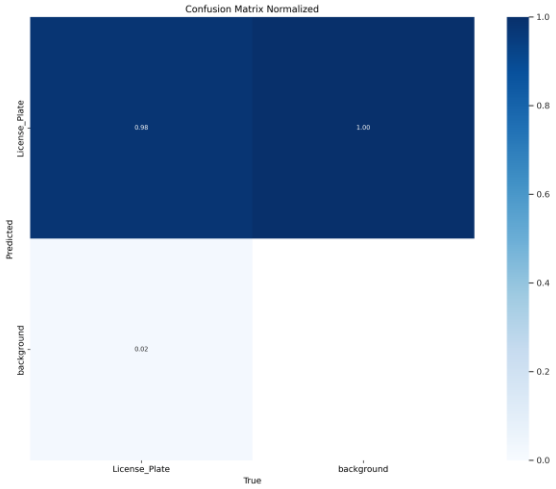


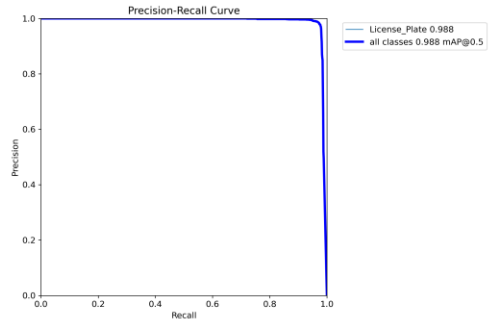












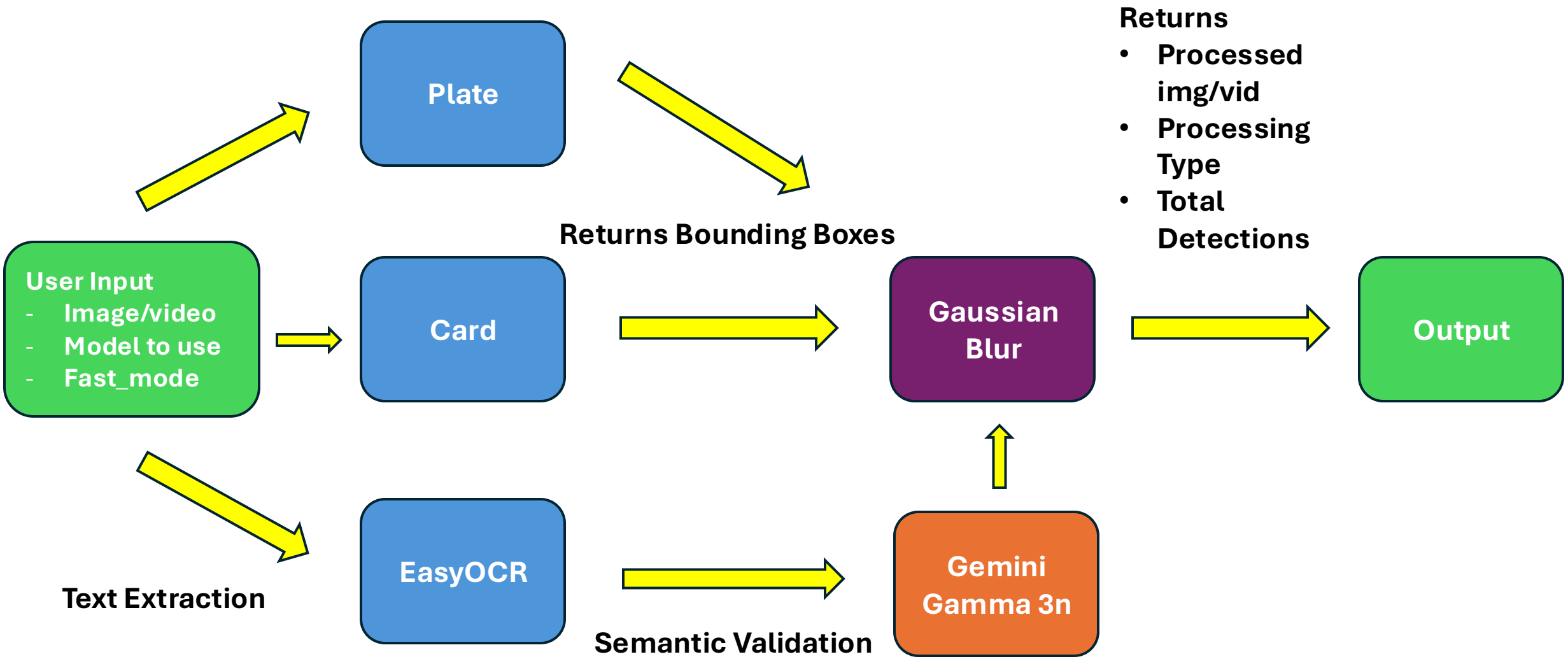
Evaluation Metrics

Evaluation Metrics

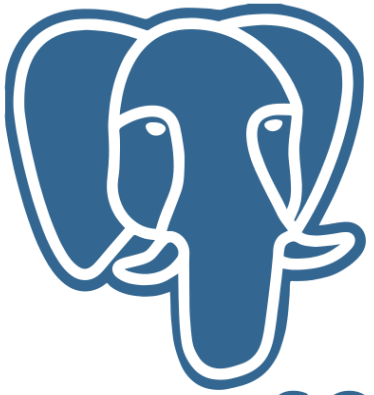
For the license plate model, training stabilized around 60 epochs. The final evaluation achieved a precision of 98.5%, a recall of 96.7%, and a mean Average Precision at IoU 0.5 of 98.9%. The stricter mAP@0.5–0.95 stood at 72.6% reflecting reason able performance under varying IoU thresholds. While precision is high, recall is the more critical number here. A recall of 96.7% means that most plates were detected and masked. Only a very small fraction might slip through without being hidden, and that is the risk we want to keep as low as possible.

The banking card model needed more training, about 100 epochs, due to the smaller data set. The results were even stronger, with precision at 99.4%, recall at 99.2%, and mAP@0.5 at 99.2%. Its stricter mAP@0.5–0.95 reached 86.7%, higher than the plate model. Once again, recall is the key metric. At 99.2%, the model missed almost no card instances. This is crucial since even a single uncovered card detail, like a number or CVV, could enable unauthorized transactions. Extra masking from false positives is acceptable, but missing a true card detail is not.

System Flowchart



Application interface here



Tools & Technologies

Conclusion

This project implemented a computer vision–based framework capable of detecting and concealing sensitive information in visual media.

High precision and recall across multiple models validate the system’s robustness and reliability.

The proposed approach demonstrates significant potential for real-world deployment in enhancing user privacy and data protection.

By integrating advanced detection with automated redaction, *Protective Vision* contributes toward developing secure, ethical, and intelligent privacy-preserving systems.

Future Work

- Extend detection beyond plates and cards to include ID cards, handwritten notes, QR codes, and official documents.
- Use lighter YOLO versions, model pruning, and compression for faster results.
- Ensure smooth real-time on-mobile processing even on low-end smartphones.
- Integrate into OLX, PakWheels, or other platforms for automatic privacy protection in user-uploaded images.
- Add features like customizable masking styles, multilingual OCR, and voice commands.
- Research to be continued

Thank You

Any Questions?