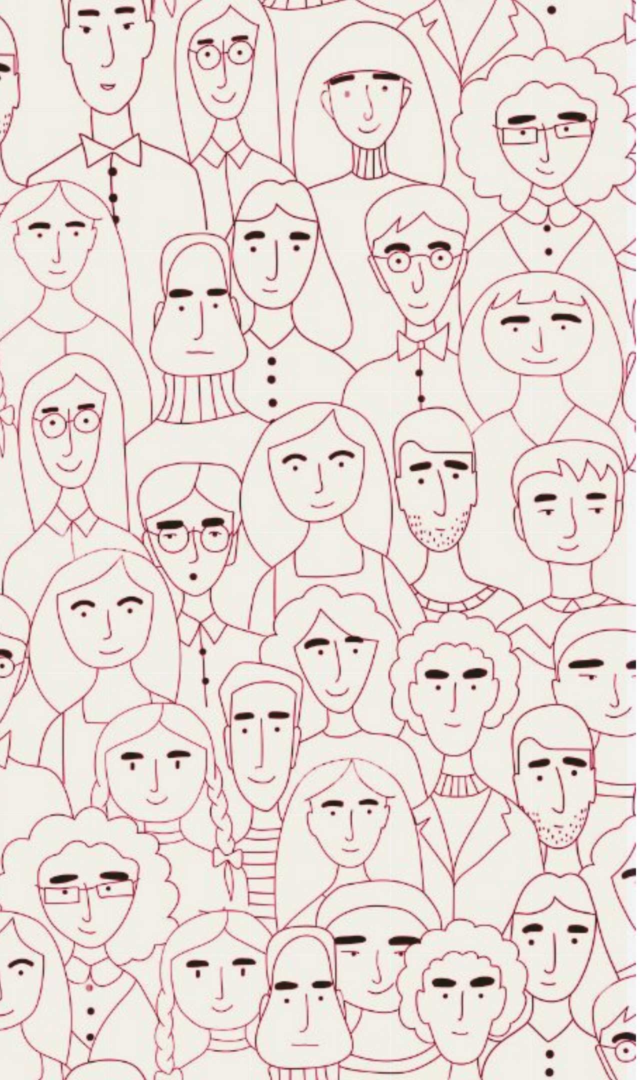
CIS 515

Computer Vision for Facial Identification at ASU

By Team 424:
Hannah Shen, Nilesh Rakhecha,
Sai Karun Reddy Gaddam, Sravya Velamuri

Arizona State University



Project Overview

Part 1 Problem definition

Proposed end to end solution

Part 2 Impact

Scope and Stakeholders

Part 3 Implementation

Part 4 Deployment (Validation, Updation, Monitoring)

Benefits, Risks & Comparisons

Part 5 CV Model

Problem Definition

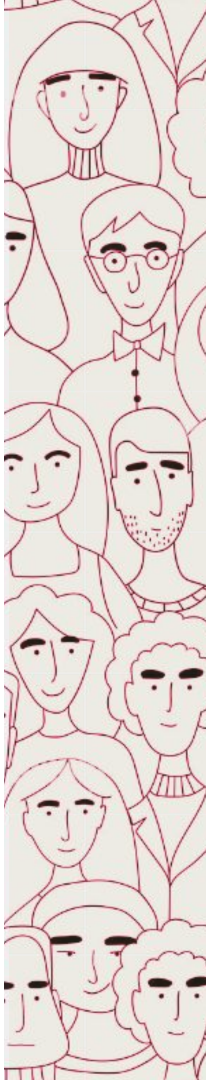
Traditional methods of attendance and RSVPing in colleges can be quite tedious, time-consuming, and prone to errors. In response, institutions have explored alternative methods such as RFID, fingerprint scanning, and passcodes. However, these methods aren't always the most satisfactory solutions, as they might not fully eliminate the challenges.

Why is it worth solving?

- With advancements in biometric technologies, facial recognition offers us a more secure and adept solution.
- To reduce the inconvenience of managing physical IDs
- Minimize security breaches
- Could help position ASU as a leader in adopting innovative security technologies in higher education

How does it concern us (team 424)?

- Our own personal experiences of misplacing the Sun Cards.
- Witnessing the hassle and hard work of faculty, TAs and Celestia during quizzes and events.
- Viability of facial recognition due to its accessibility and non-intrusiveness as a biometric feature.



Current Issues at Hand...

Vulnerability

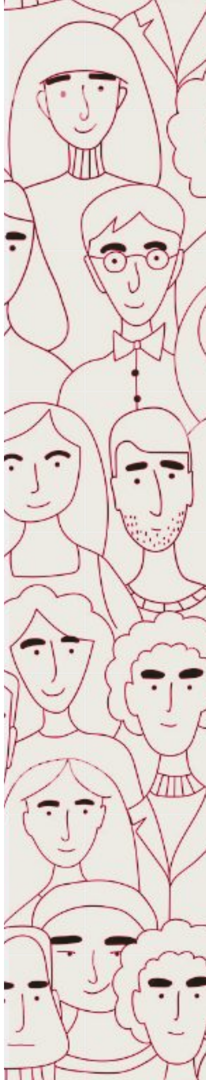
Unauthorized Access due to Sun Card being lost,stolen or shared!

Inefficiency

Attendance/RSVP/ Entry Methods are being done manually, room for errors to occur.

Risk

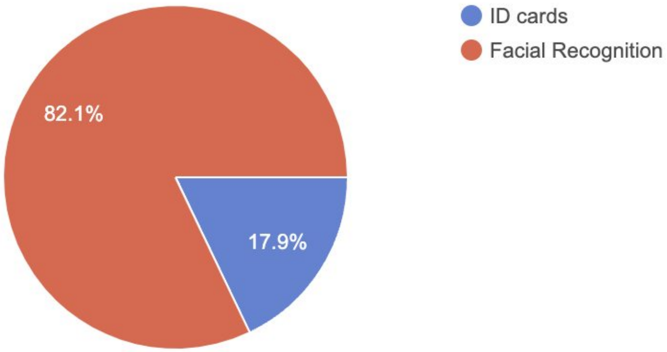
There is **potential** for **disruption** of ASU activities and a **compromise of safety**.



What do the people think?

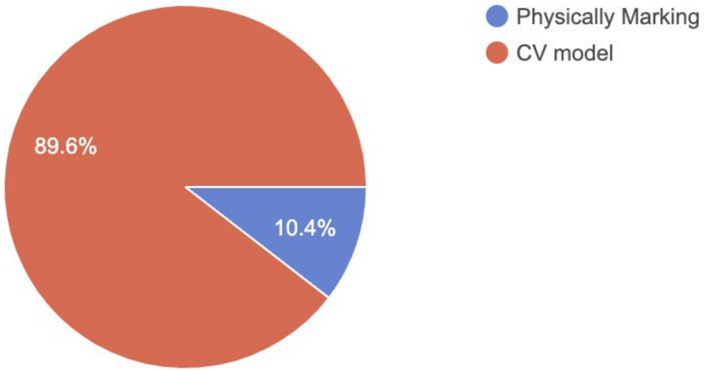
For entry into classes, team rooms, library, or other university areas, would you prefer using traditional security systems with physical ID cards or modern systems with facial recognition?

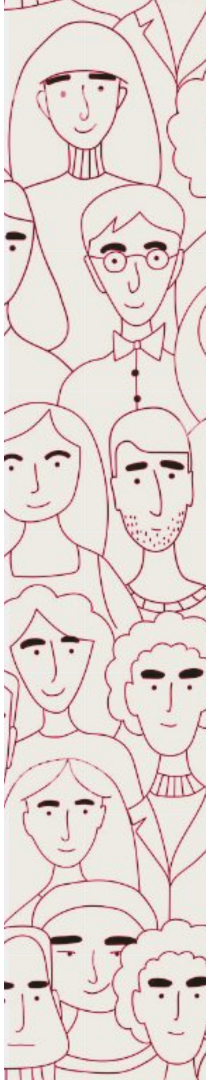
*67 Responses



When you attend an event or class, would you prefer to physically mark your attendance/check in, or have a computer vision model automatically record your presence?

*67 Responses





Proposed End to End Solution

Consent Acquisition → Image Collection and Storage → Facial Recognition at Entry Points → Attendance Marking → System Integration → Deployment

Consent Acquisition

Students are required to complete a **mandatory consent form** to authorize the use of their facial images for identification purposes.

Students who decline are redirected to an alternative identification process.

Image Collection and Storage

The **pictures** of the consenting students are taken and **uploaded on to the database**.

Then these photos are processed through our **CV model** to create **facial templates** for recognition and stored on a server.

Facial Recognition at Entry Points

Cameras installed at the **entry/exit points** of lecture halls, study rooms and event areas to **capture the faces** of students.

Then the CV model compares captured facial data with stored templates to verify identities.

Attendance Marking

Upon **successful facial match,** the system **grants entry** and **marks students as 'present**'.

Real-time attendance data, including late arrivals, is displayed on a **dashboard** accessible to professors.

System Integration

The goal is to **integrate** this with **ASU's LMS** (Canvas), to ensure **accurate** and **automated record-keeping.**

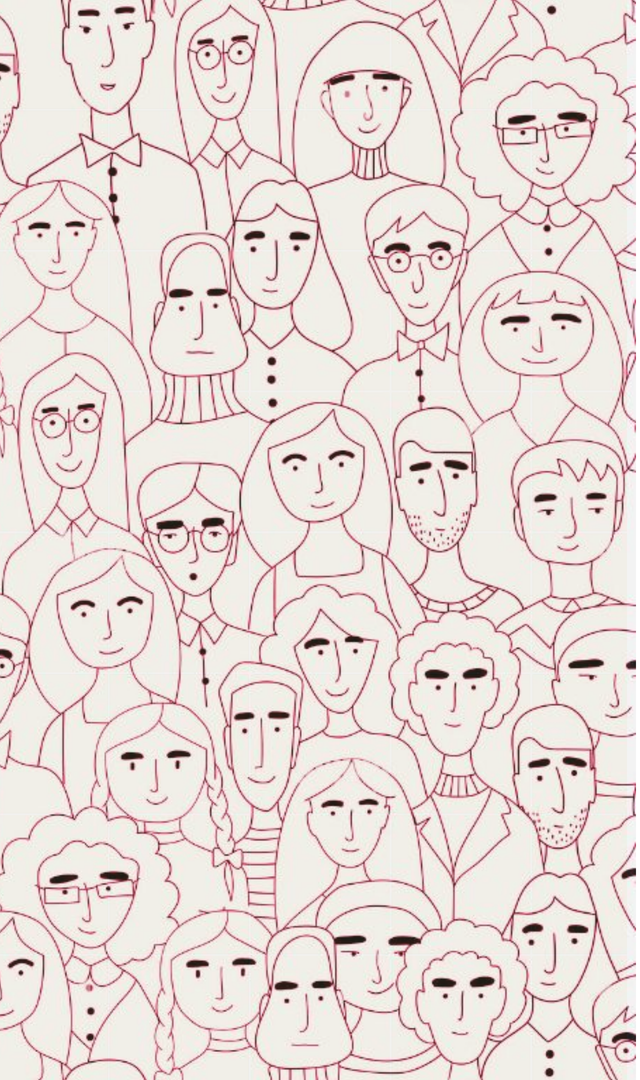
Deployment

Preparation: Perform extensive testing to ensure reliability and accuracy.

Implementation: Roll out in phases, starting with high-traffic areas and gather feedback.

Optimization: Update and refine continuously based on operational data and user suggestions.

Continuous Feedback Loop



Impact

Utility to end-user, Overall Impact, Success Metrics

Utility to End-User:

Convenience: Students and staff can access campus facilities and log attendance without the need to carry physical IDs.

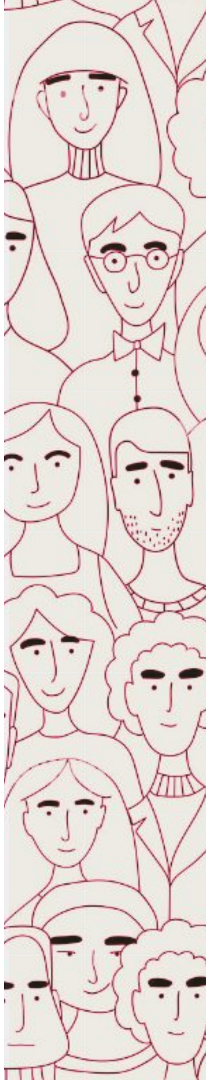
Speed: The system automates entry and attendance processes, significantly reducing the time required compared to manual checks.

Overall Impact:

➔ **Security Enhancement:** By using biometric data (facial recognition), the system minimizes the risk associated with lost, stolen, or shared ID cards.

➔ **Operational Efficiency:** Streamlining the attendance and access processes improves the overall operational efficiency of campus management.

➔ **Compliance and Record-Keeping:** Automatic attendance updates and integration with LMS ensure better compliance with institutional policies and easier record-keeping.



Cost-benefit tradeoff

Costs

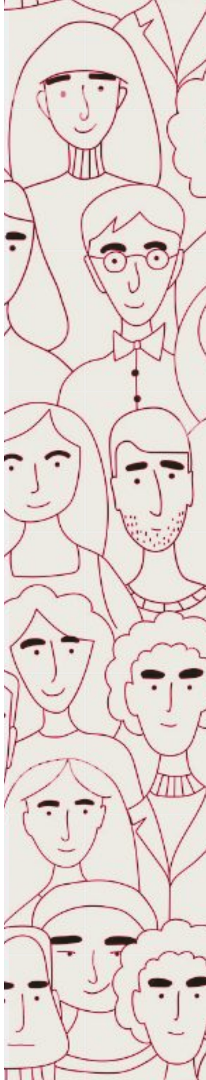
Initial Setup: Includes the cost of cameras and servers to process and store facial data.

Maintenance: Software Updates, system maintenance, additional security measures.

Benefits

Reduced Labor Costs: Minimizes the need for manual attendance taking and security checks.

Decreased Fraud: Limits cases of attendance fraud and unauthorized access.



Success Metrics

Adoption Rate
High use by students and faculty suggests ease of system adoption

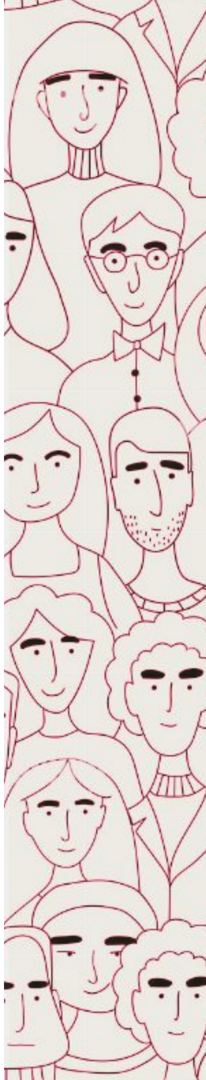
Time Saved
Quicker entries and check-ins

Unauthorised Access Reduction
Less unwanted entries

User Satisfaction
Positive survey feedback

Technical Issues Incidence
Minimal system errors

Accuracy
High Correct identification rate

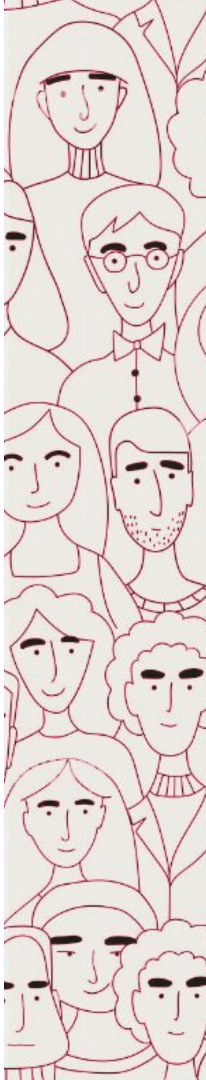


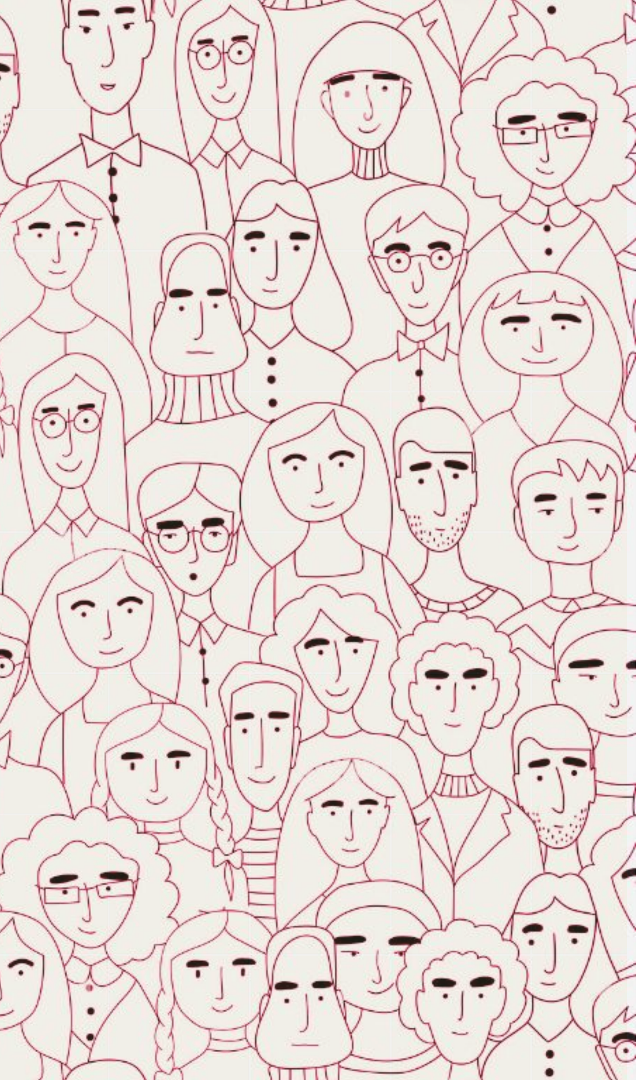
Drawing out estimates...

Based on the **outcomes** of similar solutions at **China Pharmaceutical University & Hyogo College of Medicine**, we anticipate our system here will function comparably.

We plan to conduct, the following:

- ➢ Comparative Analysis
- ➢ Cost Analysis
- ➢ Security Evaluation
- ➢ Feedback and surveys





Scope and Stakeholders

Who benefits and how?

Scope of our project

Campus wide Implementation

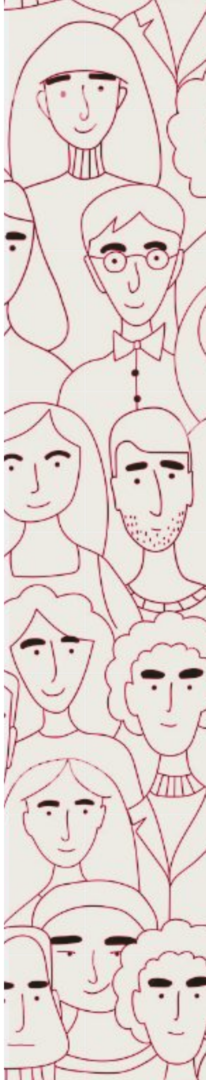
Designed to cover all access points across ASU, including lecture halls, study rooms and event venues, to streamline entry and monitoring.

Automated Attendance tracking

Integrates with ASU's LMS to automate attendance recording for classes and campus events.

Real - time Monitoring

Provides real-time updates and attendance reports to faculties.



Stakeholders

01 **ASU Administration**
Oversees system implementation and management

02 **ASU IT Department**
Handles the technical setup and maintenance

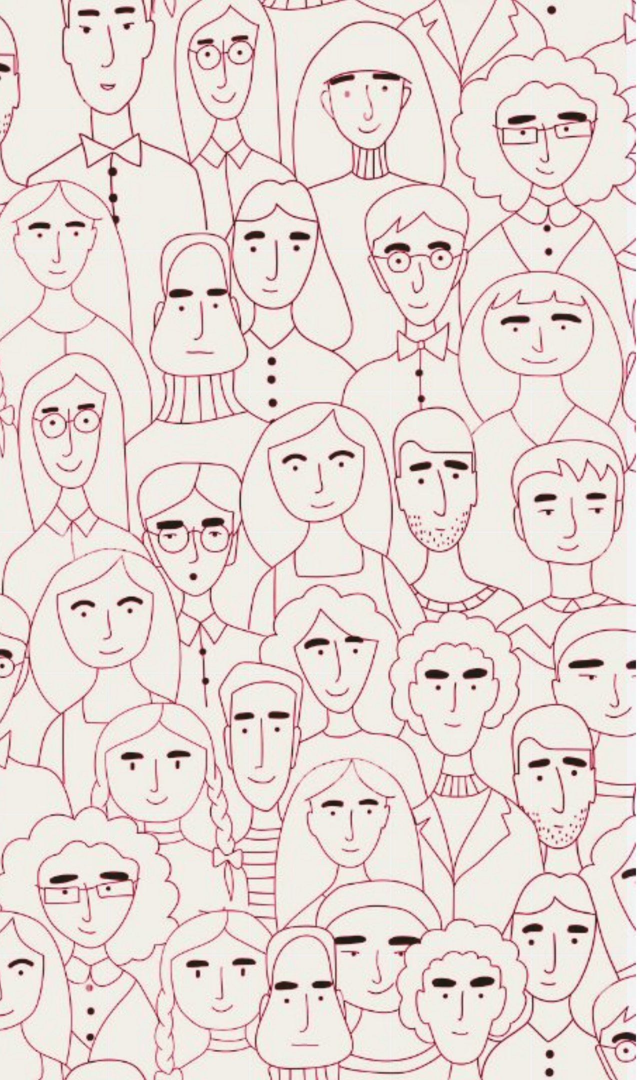
03 **Campus Security**
Manages security protocols and enhances campus safety

Beneficiaries

Student 01
Improved security and automated attendance

Faculty 02
Reliable attendance tracking and access control

Staff 03
Simplified access to work areas



Implementation

How is it going to be executed?

Estimated Resources and Cost

People

- Project Managers
- IT Security Experts
- Developers
- Maintenance Team

- **Monthly Salaries**

Data

- Facial Images of Students and Staff
- Continuous ongoing data collection process

- **Costs for taking facial images and storage costs**

Systems

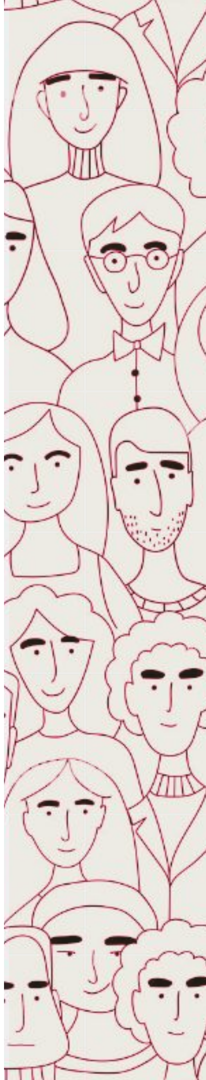
- Cameras at Entry Points
- Servers for Storage and Data processing Purposes

- **Cameras cost**
- **Installation and maintenance of cameras**
- **Server Charges**

Computational Resources

- High Performance Computing Resource

- **Cost for software development**
- **System Maintenance Cost**



Workflow

People

Training faculty, staff and IT support & Hiring additional personnel

Costs : Training & Hiring costs

Process

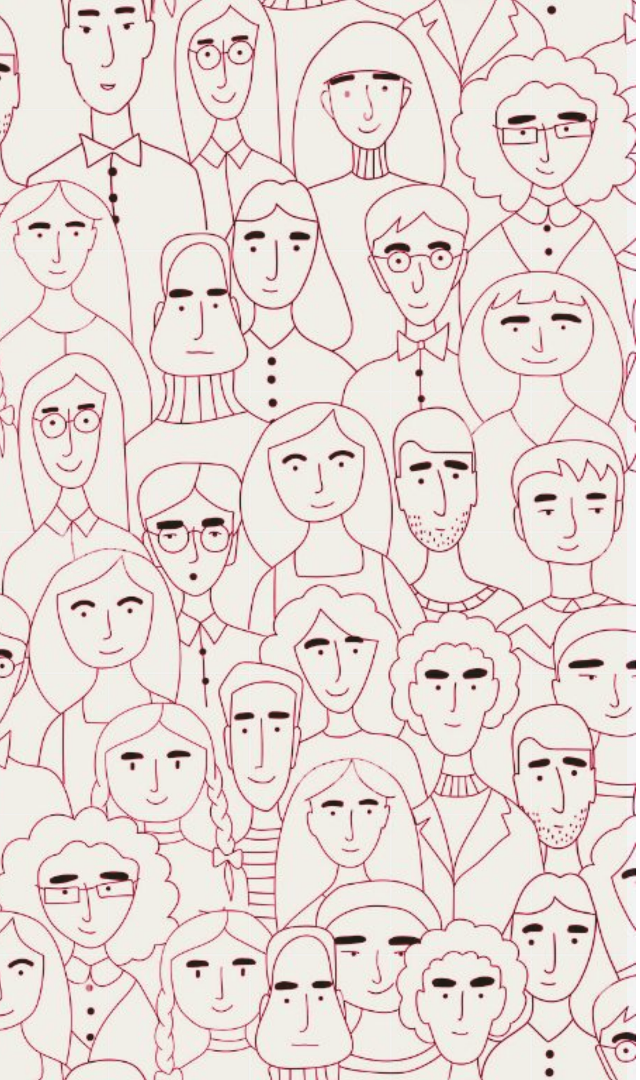
Cameras installation, data collection and Updating existing systems

Costs : Cloud service, Installation & Downtime costs

Technology

Software and hardware updates, Cyber security

Costs : Updates, Adapting to existing infrastructure & Security



Deployment

Validation and Monitoring

Pre Deployment

Pilot testing
Conduct pilot tests in controlled environments(small area) to assess the system’s accuracy and reliability.

Stress testing
Simulate peak loads to evaluate performance under stress, including rapid consecutive entries and high traffic scenarios.

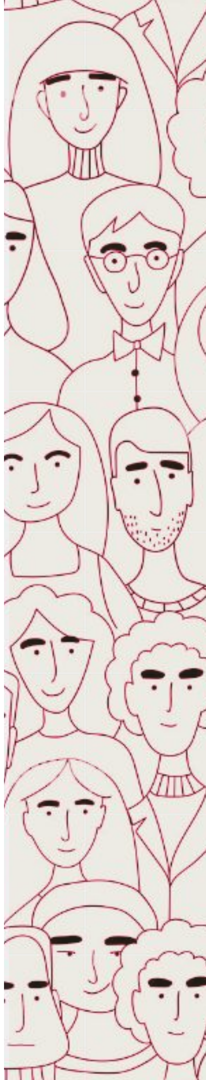
Compliance review
Ensure all system components comply with legal and privacy requirements.

Post Deployment

Performance metrics
Track system accuracy, processing speed, and gather user feedback to ensure optimal performance and satisfaction.

Security Audits
Regular security checks and vulnerability scans to identify and address potential threats

Iterative Updates
Keep the system up-to-date with the latest facial recognition advancements and user feedback



Risks and Solutions

Unfair Treatment

Facial recognition could be unfair to some groups based on how they look

Solution: Regularly evaluate and update the system to make sure it's fair and and fix any biases

Technology Failures

Entry restrictions during system breakdowns and possibility of locking people inside

Solution: A backup solution to maintain access during tech failures.

Security Risks

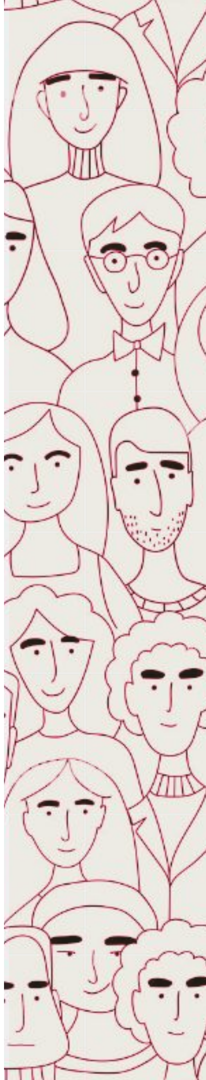
Information stored could be a target for hackers & risk of improper use of data.

Solution: Regular security updates and checks for vulnerabilities

Privacy Concerns

Students and staff might feel uncomfortable if they think they're being watched all the time.

Solution: Make sure the data is only accessible to certain people, protect it with strong security measures



Existing Implementations

China Pharmaceutical University - Uses Facial recognition for attendance and behavioral monitoring.

Hyogo College of Medicine, Japan - Uses facial recognition for student attendance.

Differences with Our System:

- **Privacy and Consent:** Our system prioritizes user consent and privacy, unlike broader surveillance models.
- **Community Feedback and Adaptation:** We develop our system with feedback from ASU's community, tailoring it to their needs.

- References
  - https://www.scmp.com/news/china/science/article/3025329/watch-and-learn-chinese-university-says-new-classroom-facial
  - https://www.campusidnews.com/japanese-college-using-facial-recognition-to-log-attendance/

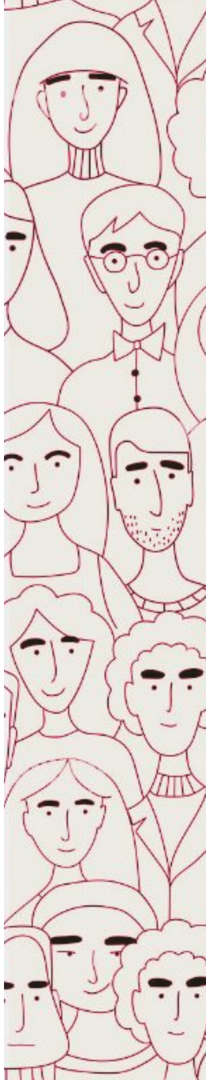
Way Forward

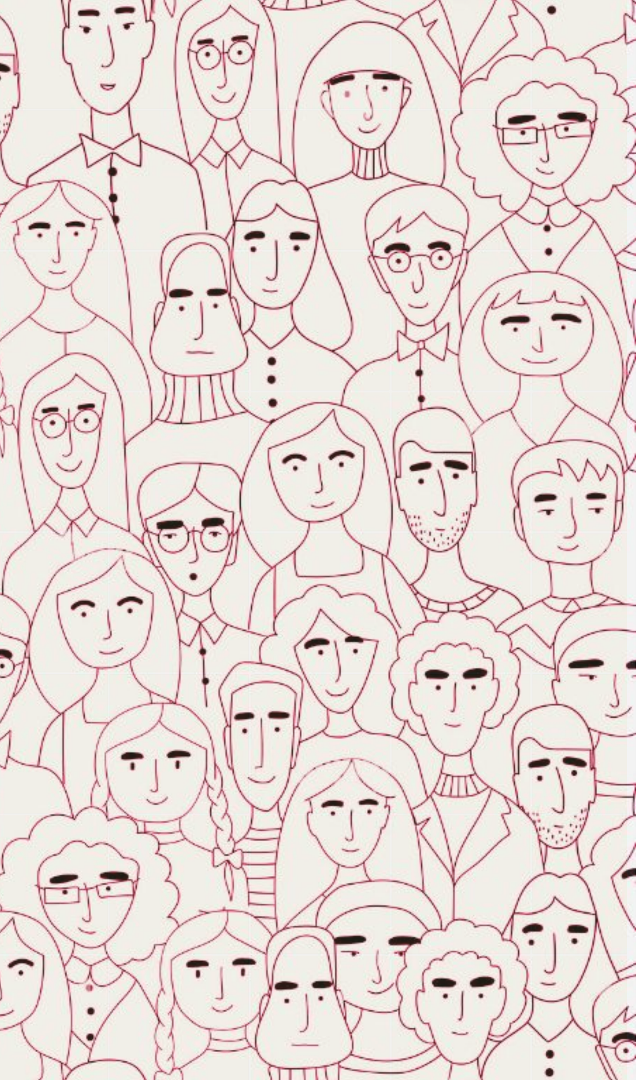
As we developed and trained the model we learnt the importance of:

- **Data Collection**
- **Parameter Count**

And also how else can the model be used in the future:

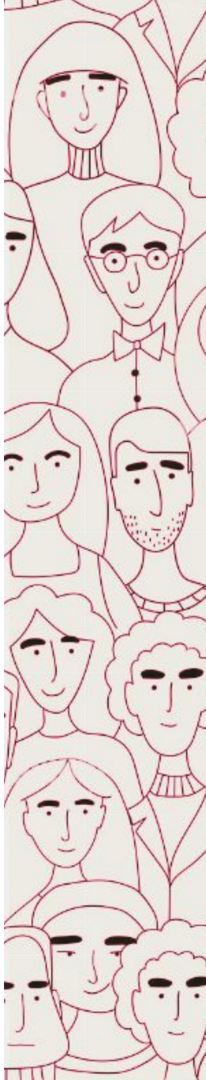
Integration with other processes	Adapt the system for use in different contexts within the university, such as for fines in no wheel zones, security in library, MU, and other areas.
Behavioural Assessment	The model can be further developed to be incorporated along with in-class cameras and lockdown browser to analyse student behavior.





CV Model

Proof Of Concept Demo (pt.1 - Shoot)

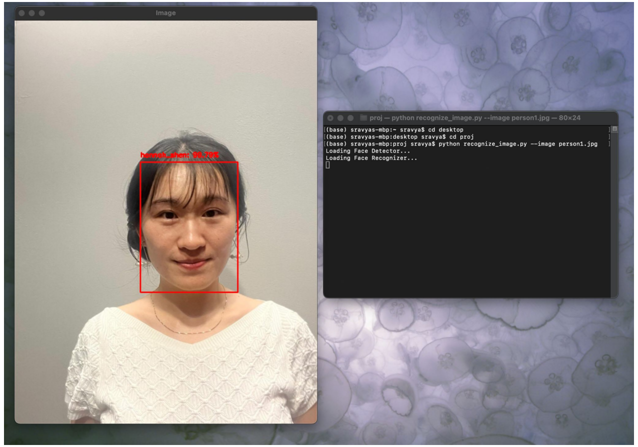


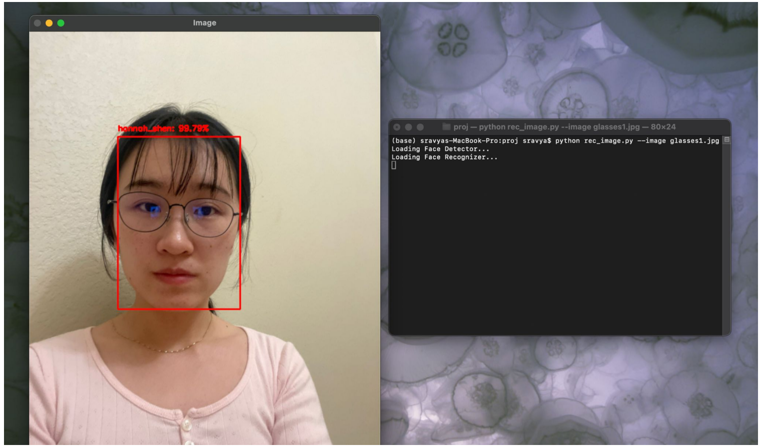
We have used the Haar classifier[1] to identify the face, it works in the following way:

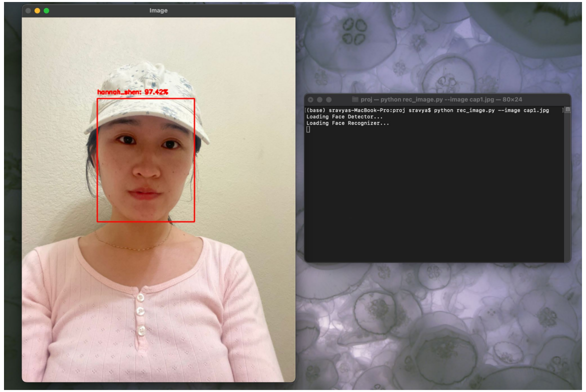
1. Identifies objects like faces in images by comparing pixel intensities.
2. Uses rectangle-based patterns to distinguish objects from their surroundings
3. The cascading function eliminates irrelevant areas quickly,so apt for live applications.

Reference: 1- https://github.com/topics/haar-cascade-classifier

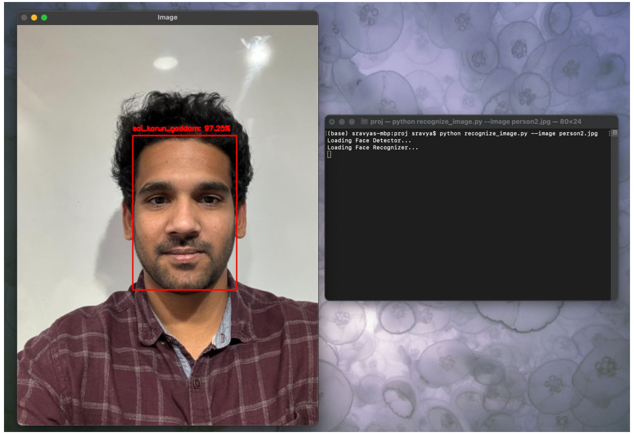
Proof Of Concept Demo (pt. 2 - Image Recognition)

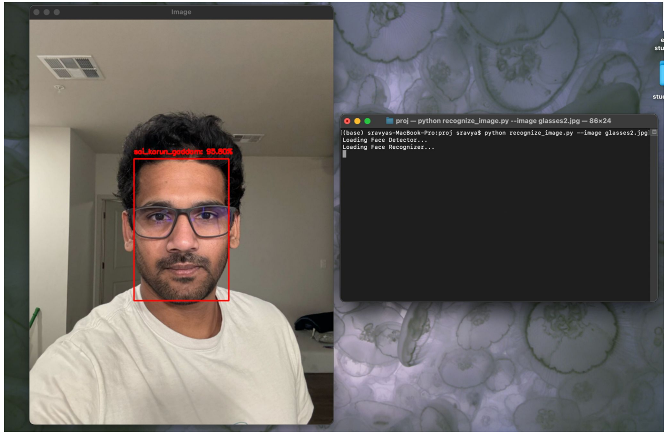


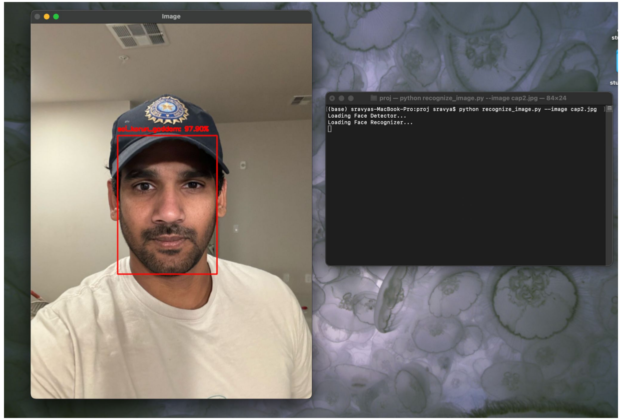


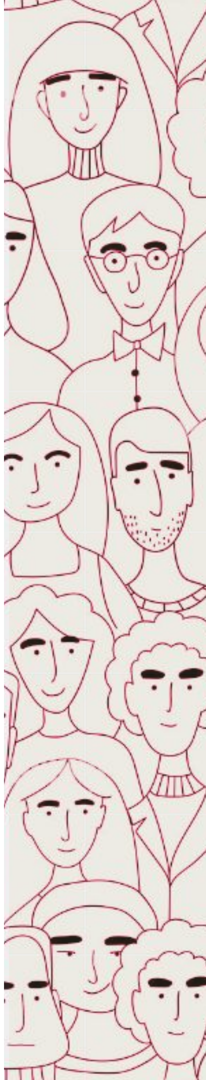


Proof Of Concept Demo (pt. 2 - Image Recognition)

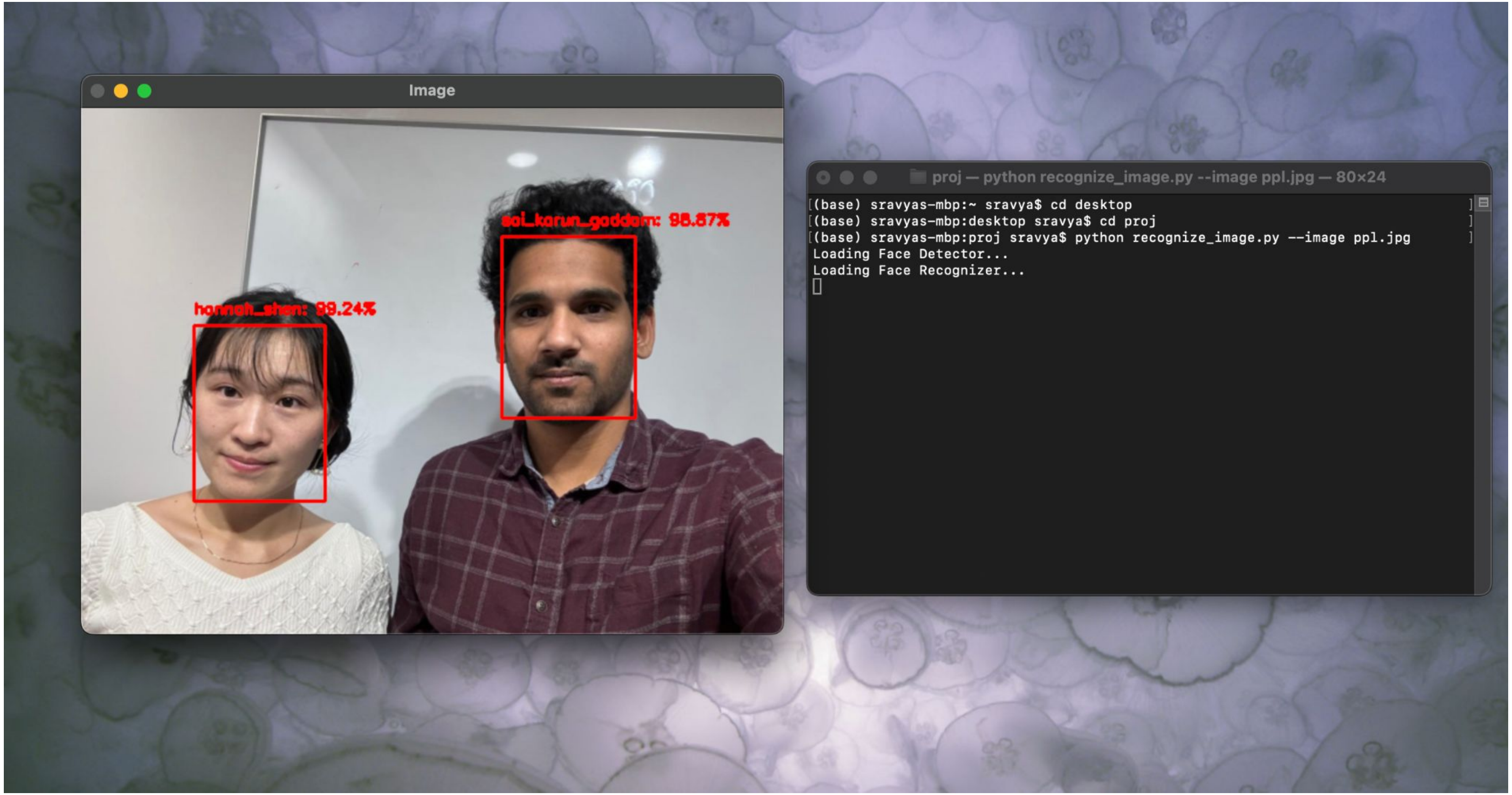


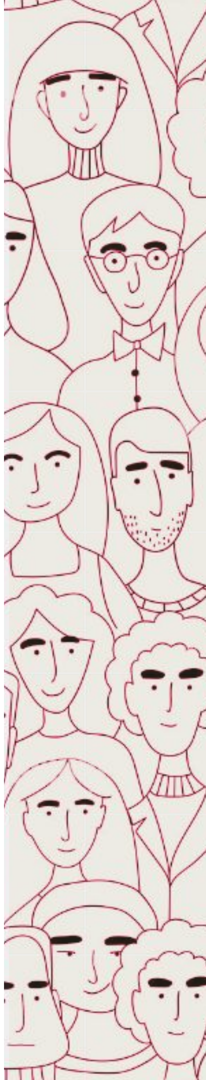






Proof Of Concept Demo (pt. 2 - Image Recognition)





Proof Of Concept Demo (pt. 3 - Live Video Recognition)

About our CV model

Model Summary: We employed **transfer learning** on our chosen open source model (**ResNet-50**[1]), a CNN model initialized with weights pre-trained on the **VGGFace dataset**[2], a large-scale face recognition dataset.

During transfer learning, the pre-trained ResNet-50 model[1] serves as a **feature extractor**, capturing facial features from our input images. These learned features are then fine-tuned for the specific face classification task at hand.

1. **Preprocessing**: Images are preprocessed to ensure that the input images are suitable for the VGGFace model.

2. **Training:** The model is trained initially for 50 epochs using the training and testing data generators, using the Adam optimizer with a learning rate of $1e^{-4}$

3. **Fine-Tuning:** After training, the pre-trained layers are unfrozen, and the model is fine-tuned for an additional 10 epochs using a lower learning rate ($1e^{-5}$).

A Deeper Look into ResNet-50 and VGGFace Dataset

ResNet-50[1] is a **50-layer convolutional neural network** (48 convolutional layers, one MaxPool layer, and one average pool layer).

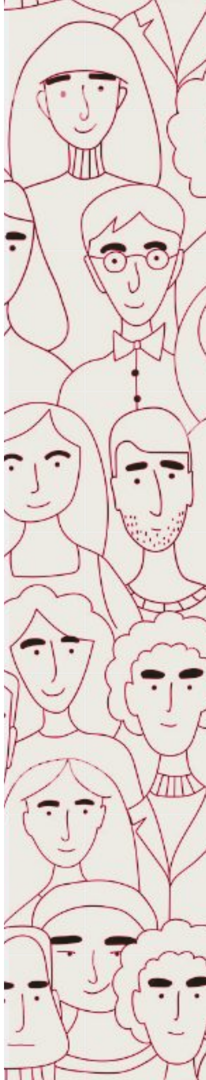
The VGG Face dataset[2] is face identity recognition dataset that consists of **2,622 identities**. It contains over **2.6 million images**.The VGG Face dataset was created to provide access to biometric data to researchers working on face recognition technologies.

References:
[1] - https://pytorch.org/hub/nvidia_deeplearningexamples_resnet50/
[2]- https://exposing.ai/vgg_face/

Our Dataset:

Consists **105 celebrities** and **17534 face pics**, that have been collected from Pinterest and cropped, along with our mini data set consisting of **3 of our team members** (Sai, Sravya, Hannah) and **288 face pics**. Total face pics = 17,822.

Dataset Link: https://www.kaggle.com/datasets/hereisburak/pins-face-recognition



Our Training Process

1. During each epoch, the model iterates over the training dataset in batches provided by the train_generator.

2. For each batch, the model computes predictions for the input data, calculates the loss (difference between predictions and actual targets), and updates parameters to minimize this loss.

3. After processing all batches in the training dataset, the model evaluates its performance on the validation dataset using the test_generator.

4. This process is repeated for the specified number of epochs (50).

Our Fine Tuning Process

1. The pre-trained layers of the model are unfrozen

2. A smaller learning rate is employed to mitigate the risk of overfitting

3. During each epoch, the model iterates over the training dataset in batches provided by the train_generator.

4. For each batch, the model computes predictions for the input data and calculates the loss

5. After processing all batches in the training dataset, the model evaluates its performance on the validation dataset using the test_generator.

6. This process is repeated for the specified number of of fine-tuning epochs (10).

A further look into Model Evaluation

The **model's performance** is evaluated using 2 primary metrics:

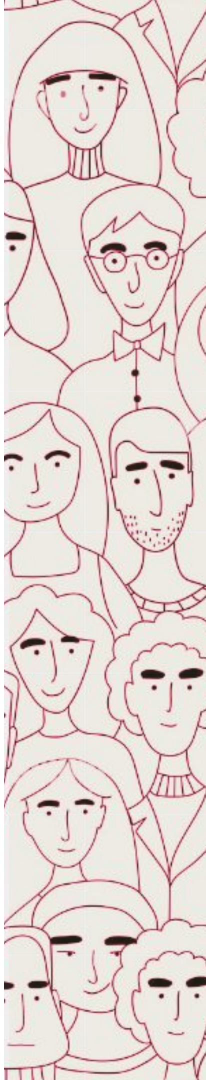
- **LOSS** (categorical cross-entropy is used as our loss function)
- **ACCURACY SCORE** (the percentage of correctly classified images)

Internal model validation uses the above metrics to monitor the model's performance over epochs and make model adjustments accordingly, such as:

- **Adjusting hyperparameters**
- **Terminating training** if overfitting is observed.

External Model Validation:

Pre-deployment: using labeled face image datasets to assess its performance across diverse demographics and conditions.
Post-deployment: collecting feedback from users after real time usage



Outcome-Action Pairings

Reevaluate classification threshold
feature engineering, error analysis

True Positive (TP) Correctly identifies student	**False Positive (FP)** Incorrectly identified a non-student as a student
False Negative (FN) Incorrectly identifies student as unknown	**True Negative (TN)** Correctly identifies non-students as unknown

Reevaluate model parameters / dataset
Look into hyper parameter tuning

Bias Issues and its Mitigation

->Data Collection Bias: Limited dataset diversity can lead to bias. For instance, The VGGFace dataset despite being widely used, falls short of representing the diversity of the global population. As noted in Oxford's research[3], its composition may harbor biases.

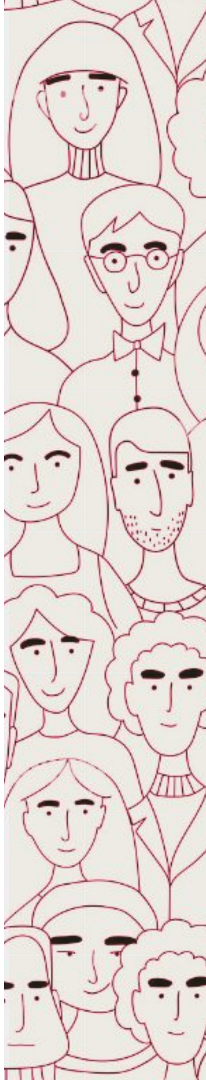
Solution: Gather a dataset that accurately mirror the demographics of our target population (students of ASU) and is balanced through all classes (ethnicities).

->Labeling Bias: Inconsistent labeling of facial features/expressions due to subjectivity can lead to inaccuracies

Solution: Ensure consistent, unbiased labeling (multiple annotators and clear guidelines)

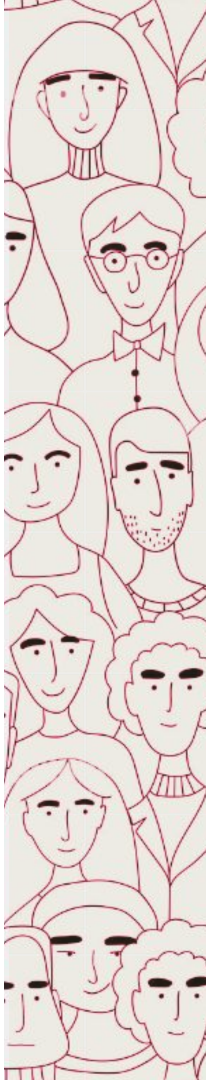
->Evaluation Bias: Accuracy Score may not be suitable for imbalanced classes, as thy favour certain group over others.
Solution: Use fairness-aware metrics like demographic parity for equitable evaluation.

References:
[3]- https://www.robots.ox.ac.uk/~vgg/data/vgg_face/



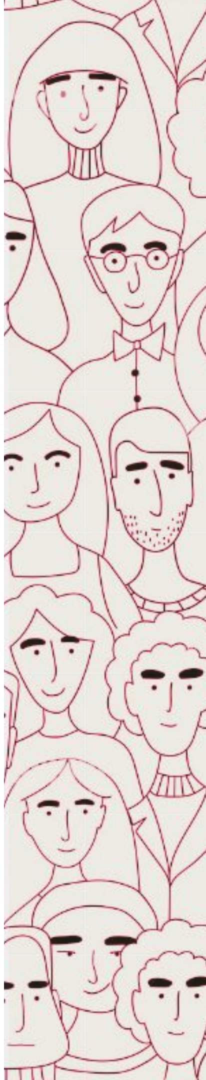
Limitations of Our Model

- **Data Quality:** Limited / biased datasets can be detrimental, as discussed earlier !
- **Lighting Variations:** Non-uniform lighting, shadows, and glare can affect facial features, leading to reduced accuracy under varying lighting conditions.
- **Facial Hair Variability:** Changes in facial hair, such as growing a beard or shaving it, introduce variability in facial appearance, impacting recognition performance.
- **Accessories:** compromisation via accessories like masks and sunglasses, which obscure key facial features (nose+mouth/eyes).
- **Weight Changes:** Significant weight gain /loss could alter facial features, so it could be challenging to accurately match faces across different weights.



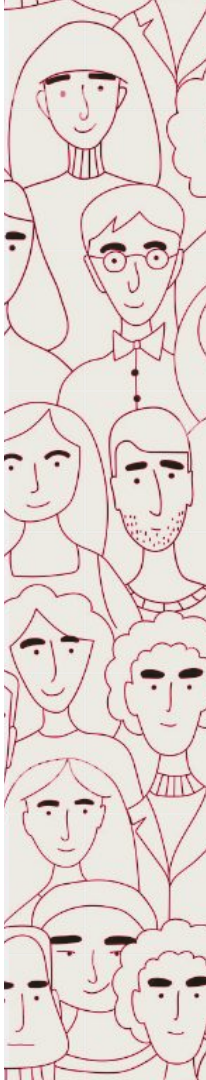
Future Implementation of Our Model (Enhancements)

- **Hyperparameter Tuning:** Fine-tune hyperparameters such as learning rate, batch size, dropout rate, and optimizer choice.
- **Data Augmentation:** Explore more sophisticated domain specific data augmentation techniques.
- **Advanced Architectures:** Consider using advanced architectures tailored for specific tasks, such as attention mechanisms for focusing on relevant parts of the image or capsule networks for handling hierarchical relationships in features.
- **Class Imbalance Handling:** employ balancing techniques that ensure the model learns equally from all classes.
- **Model Compression:** Investigate techniques for model compression and optimization, such as quantization, pruning, or knowledge distillation, to reduce model size and improve inference speed, which is crucial for deployment on resource-constrained devices.

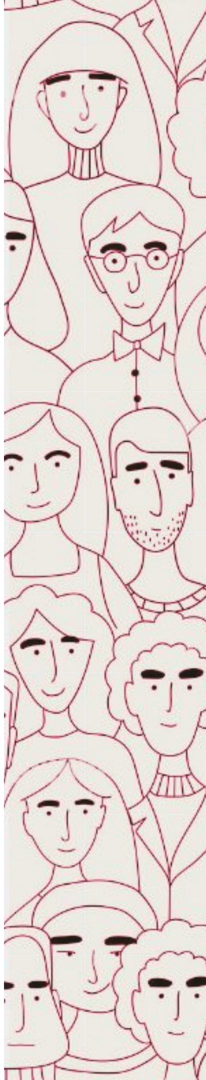


Future Implementation of Our Model (Enhancements)

- **Hyperparameter Tuning:** Fine-tune hyperparameters such as learning rate, batch size, dropout rate, and optimizer choice.
- **Data Augmentation:** Explore more sophisticated domain specific data augmentation techniques.
- **Advanced Architectures:** Consider using advanced architectures tailored for specific tasks, such as attention mechanisms for focusing on relevant parts of the image or capsule networks for handling hierarchical relationships in features.
- **Class Imbalance Handling:** employ balancing techniques that ensure the model learns equally from all classes.
- **Model Compression:** Investigate techniques for model compression and optimization, such as quantization, pruning, or knowledge distillation, to reduce model size and improve inference speed, which is crucial for deployment on resource-constrained devices.

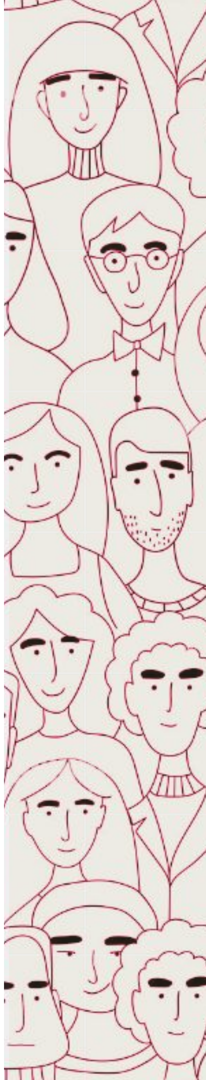


THANK YOU!



Task Ownership

Survey	Sai Karun Reddy Gaddam
Background research	Hannah Shen
Problem definition	Sravya Velamuri
Proposed End to end solution	Everyone
Impact	Nilesh Rakhecha
Scope and Stakeholders	Sai Karun Reddy Gaddam
Estimated costs and workflow	Nilesh Rakhecha
Model Selection & Finding Set	Sravya Velamuri
Creating our dataset	Hannah Shen
Shooting and face Detection	Sai Karun Reddy Gaddam
Model Develop and train	Sravya Velamuri
Validation	Nilesh Rakhecha
2D image recognition	Hannah Shen
Live Video recognition	Sravya Velamuri
Slides	Everyone



Paper Reference

- https://www.ijert.org/volume-09-issue-05-may-2020

Code Reference

- https://github.com/aakashjhawar/face-recognition-using-deep-learning
- https://github.com/Srikeshram/Celebrity-Face-Recognition