**Coursework 1: Age Estimation and Gender Classification Report – Plan**

**1. Introduction**

* **Problem statement**: Why is age estimation and gender classification important? (Security, retail, social media, healthcare)
* **Dataset description**: Using a **5,000-image subset of UTKFace**, which includes labeled facial images (not the full ImageNet dataset).
* **Objective**: Train and evaluate two CNN models:
  + **Model A**: Custom CNN trained from scratch.
  + **Model B**: Pre-trained MobileNetV2 fine-tuned on our dataset.
* **Evaluation metrics**:
  + **Age Estimation**: Mean Absolute Error (MAE).
  + **Gender Classification**: Accuracy.
* **Report structure**: Preprocessing, model architectures, training process, results, and comparison.

**2. Data Preprocessing (4 marks)**

* **Dataset Characteristics**:
  + **Images are labeled with age and gender.**
  + **Challenges**: Variations in pose, lighting, expressions, and image quality.
* **Steps Taken**:
  + **Sample visualization**: Displayed a few images with age and gender labels.
  + **Normalization**: Pixel values scaled to [0,1] to stabilize training.
  + **Data Augmentation**: Applied transformations to prevent overfitting:
    - Rotation, flipping, zooming.
    - Justification: Augmentation increases generalization and robustness.
  + **Splitting dataset**:
    - **80% Training, 20% Validation**.
    - **Why?** Ensures model learns effectively while leaving data for evaluation.

**3. Model A: Custom CNN (44 marks)**

**3.1 Model Construction (12 marks)**

* **Architecture**:
  + Input shape: **(128, 128, 3)**
  + **4 Convolutional Layers**:
    - Filters: **32, 64, 128, 256** (increasing complexity).
    - **ReLU activation**: Avoids vanishing gradients.
    - **MaxPooling**: Reduces dimensionality and computation.
  + **Flatten Layer**: Converts feature maps into a single vector.
  + **Fully Connected Layers**:
    - **Dense (256 neurons)** – Extracts high-level features.
    - **Dropout (0.3)** – Prevents overfitting.
  + **Output Layers**:
    - **Age prediction**: Linear activation (MAE loss).
    - **Gender classification**: Sigmoid activation (Binary Cross-Entropy loss).
* **Why this architecture?**
  + Deep enough to extract meaningful features.
  + Small enough to avoid overfitting with limited data.

**3.2 Effective Training (6 marks)**

* **Hyperparameters**:
  + **Optimizer**: Adam (adaptive learning rate).
  + **Learning Rate**: 0.001 (balance between speed and stability).
  + **Batch Size**: 32 (memory efficiency vs. model stability).
  + **Epochs**: 20 (enough for convergence, avoiding overfitting).
* **Learning Curves**:
  + Monitor training & validation loss.
  + Identify underfitting (high loss) or overfitting (low training loss, high validation loss).

**3.3 Model Explanation (8 marks)**

* **Discussion of architecture choice**:
  + Why **4 convolutional layers** instead of more/fewer?
  + Why **batch normalization wasn’t used**?
  + Why **dropout rate = 0.3**?
* **Training methodology**:
  + How augmentation impacted results.
  + Early stopping considerations.

**3.4 Performance (18 marks)**

* **Evaluation Metrics**:
  + **Age Estimation MAE**: TBD (Lower is better).
  + **Gender Classification Accuracy**: TBD (Higher is better).
* **Discussion of Results**:
  + How well does the model generalize?
  + Any **biases observed**? (e.g., age mispredictions in certain groups).

**4. Model B: Pre-Trained CNN (MobileNetV2) (44 marks)**

**4.1 Model Construction (12 marks)**

* **Why MobileNetV2?**
  + **Pre-trained on ImageNet**: Strong feature extraction.
  + **Efficient**: Depthwise separable convolutions reduce computation.
* **Modifications for our dataset**:
  + **Remove the original fully connected layers**.
  + **Add new layers**:
    - **GlobalAveragePooling2D** (reduces feature maps).
    - **Dense (128 neurons, ReLU activation)**.
    - **Dropout (0.3)**.
    - **Two output layers** (same as Model A).

**4.2 Effective Training (6 marks)**

* **Transfer Learning Approach**:
  + **Freeze early layers** (retain generic feature extraction).
  + **Unfreeze last few layers** (allow adaptation to our dataset).
* **Hyperparameters**:
  + **Lower learning rate (0.001)** to avoid damaging pre-trained weights.
  + **Batch Size: 32**.
  + **Epochs: 20**.

**4.3 Model Explanation (8 marks)**

* **Why fine-tuning instead of feature extraction only?**
* **What features does MobileNetV2 learn?**
* **Impact of freezing vs. unfreezing layers**.

**4.4 Performance (18 marks)**

* **Evaluation Metrics**:
  + **Age Estimation MAE**: TBD.
  + **Gender Classification Accuracy**: TBD.
* **Comparison with Model A**:
  + Did transfer learning improve accuracy?
  + Did MobileNetV2 outperform the custom CNN?
  + Limitations: Is the model biased towards certain age/gender groups?

**5. Summary & Discussion (8 marks)**

**5.1 Comparison of Models**

| **Aspect** | **Model A (Custom CNN)** | **Model B (Pre-Trained CNN)** |
| --- | --- | --- |
| Architecture | 4 Conv layers + Dense | MobileNetV2 + Custom FC |
| Training Time | Longer | Shorter due to pre-trained features |
| Age Estimation (MAE) | TBD | TBD |
| Gender Classification Accuracy | TBD | TBD |
| Overfitting Handling | Dropout layers | Transfer learning & fine-tuning |

**5.2 Strengths & Limitations**

* **Model A**:
  + ✅ Full control over architecture.
  + ❌ Requires more training time.
* **Model B**:
  + ✅ Leverages pre-trained knowledge.
  + ❌ May not generalize well to very different datasets.

**5.3 Real-World Applications**

* How could this technology be applied practically?
  + **Retail**: Targeted advertising based on age/gender.
  + **Healthcare**: Monitoring age-related conditions.
  + **Security**: ID verification systems.

**5.4 Future Work**

* **Increase dataset size** (More training data = better generalization).
* **Test different CNN architectures** (ResNet, EfficientNet).
* **Fine-tune hyperparameters** for better optimization.

**6. Conclusion**

* **Recap of findings**:
  + Model A provides flexibility but takes longer to train.
  + Model B is efficient but may need more fine-tuning.
* **Key takeaway**: Transfer learning is a powerful tool, but custom models can still perform well with careful optimization.

**7. References**

* Cite papers/tutorials used.

**8. Model Links**

* **Model A (Custom CNN):** [Insert Link]
* **Model B (Pre-Trained CNN):** [Insert Link]

This plan **includes more points than needed**, so you can pick and choose while writing on the train. Let me know if you need refinements! 🚀

# REPORT

## Introduction

Age and gender estimation are very key challenges within computer vision, Having applications in security, social media and retail. This coursework focuses on implementing CNNs to address these tasks using a subset of the UTKFace dataset containing 5000 labelled images of faces.

Our dataset includes a diverse set of age groups allowing for models to generalise across different demographics. The goal is to train and evaluate the performance of two CNN models:

* Model A. Our own custom-built CNN trained from scratch, designed specifically to extract meaningful features while making efforts to avoid overfitting
* Model B. A fine-tuned pre-trained CNN that utilises transfer learning for improved accuracy and efficiency.

Both models are assessed on their ability to predict age and classify gender. This report details the preprocessing steps, model architecture, training methodologies and evaluation results. A comparative analysis highlights the variations in performance of both approaches, offering insights into their applicability into real world situations and areas in need of further improvements.