Université Concordia University

Age Detection by Image Classification Using Convolutional Neural Networks

Group T

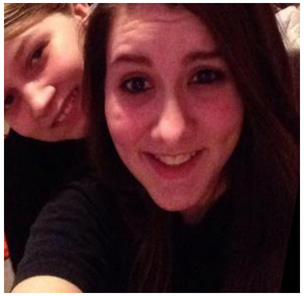
- **Mohnish Mirchandani – 40301836**
- **Santhosh Dayakar - 40302539**
- **Shashwat Shah - 40311188**
- **Vatsal Dadia - 40305289**

Introduction

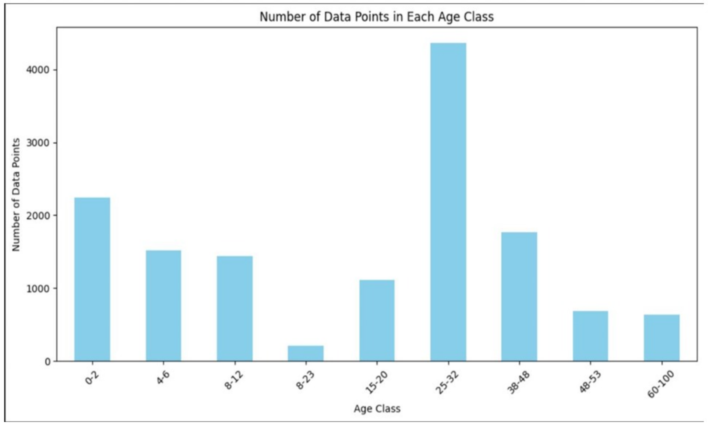
- **Importance and Use Cases of Age Detection**
  - Electronic Identity Verification
  - Healthcare and Military Applications
  - Australia's recent TDIF framework and changes.
- **Challenges in Age Classification**
  - Overlap in visual markers
  - Varying and Complex features
  - Constraints of Traditional Methods.
- **Advancements with CNNs**
  - Convolutional Neural Networks as a solution to these challenges.

Datasets

- **UTKFace Dataset**
  - Comprehensive dataset with 20,000 images "in the wild"
  - Divided into 15 age group classes.
  - Balanced Dataset, such that images across all age classes are covered



- **Adience Dataset**
  - 19,400 images captured in natural settings, such as unfiltered photos from platforms like Flickr.
  - Divided into 8 age groups: (0–2), (4–6), (8–12), (15–20), (25–32), (38–48), (48–53), and (60–100).



- **FacialAge Dataset**
  - 9,778 images, organized into 4 broad age groups: (0–7), (7–25), (25–49), and (49–110).
  - Generated and collected from sources like WIKI_ART, offering a mix of real-world and generated data.
  - Leans towards younger age groups.

CNN Architectures

VGGNet

- VGGNet-16 is a deep CNN known for its simple, uniform structure
- Comprises of 3×3 convolutional layers with ReLU activations and max-pooling layers.
- No Information loss due to smaller kernel and enables hierarchical features.

DenseNet

- DenseNet-121 employs densely connected blocks, where each layer receives inputs from all preceding layers
- Feature Map Concatenation and transitive layers
- Features from all previous layers are utilized allowing fine grained detail utilization

EfficientNet

- Advancement on Mobilenet; Uses Depthwise Seperable Convolutions with squeeze excitation block
- Computationally efficient
- Uses Global Pooling to enable information capture of wrinkles, skin textures etc.

Methodology: Preprocessing & Training

Data Preprocessing

- **Alignment and Cropping**
  - Resizing images to 224x224 due to model requirements.
- **Normalization**
  - According to IMAGENET standards
- **Augmentation**
  - Random rotations, flipping, color jittering, and brightness adjustments

Training Setup

- **Learning Rate Tuning**
  - Adam optimizer with learning rate of 0.001 across most instances
- **Batch Size and Epochs**
  - Models are trained using a batch size of 32 and for 5, 10 and 20 epochs. Early stopping is implemented to halt training when validation loss stops improving

Training and Evaluation

Training Process

- **Loss Function**: Cross-entropy loss is used comparing predicted probabilities with true labels.
- **Optimizer**: Adam optimizer
- **Learning Rate: Dynamic**
  - 0.001 across most
  - Step scheduling with decrease in learning rate of 0.1

Evaluation Metrics

- **Accuracy**: Minimum – VGGNet , Maximum – EfficientNet.
- **Classification Report**: Precision, Recall, F1 Score

Visualization Tools

- **t-SNE (t-Distributed Stochastic Neighbor Embedding)**:

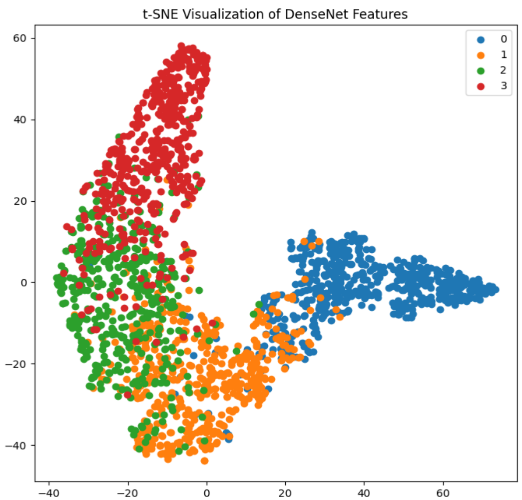
```
Test Accuracy: 81.34%
Confusion Matrix:
 [[490  36   2   0]
 [ 33 415  39   3]
 [  2 120 285  47]
 [  1   9  73 401]]

Precision per class: [0.93155894 0.71551724 0.71428571 0.88913525]
Recall per class: [0.9280303  0.84693878 0.6277533  0.8285124 ]
F1-score per class: [0.92979127 0.77570093 0.66822978 0.85775401]
Samples per class: [528 490 454 484]
```

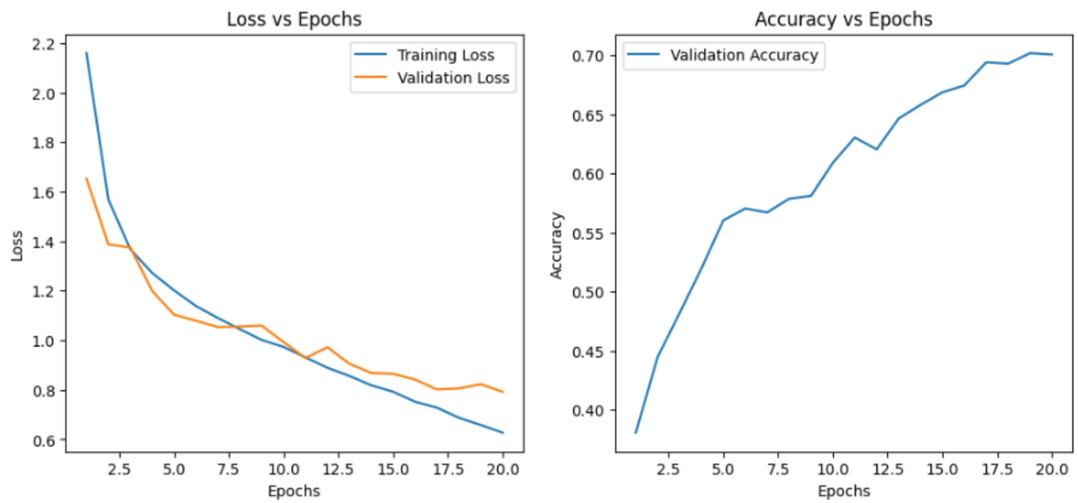
Result Overview

Model	Parameters	Test Accuracy: Facial Age (10 epochs)	Test Accuracy: UTKFaces (20 epochs)	Test Accuracy: Adience (5 epochs)
VGGNet16	~ 138M	81.34%	54.17%	60.20%
DenseNet121	~ 8M	80.57%	73.00%	70.81%
EfficientNetB0	~ 5.3M	80.14%	70.49%	76.94%
ResNet18 (Transfer Learning)	~ 11.6M	-	-	51.11%
ShuffleNet (Transfer Learning)	~ 2.3M	64.04%	-	-

Result Overview



TSNE on EfficientNet trained on FacialAge



Loss vs Epoch Curve with Accuracy: EfficientNet trained on UTKFaces

Conclusion

Conclusion

- **EfficientNet performs better**:
  - EfficientNet outperforms other models, striking the best balance between computational efficiency and accuracy.
- **VGGNet needed improvements**:
  - VGGNet consistently underperforms as compared to other models across all three datasets by significant measure.
- **Generalizability is difficult**:
  - Both models underperformed when it came to transfer learning from IMAGENET dataset.

References

- [1] Eran Eidinger, Roee Enbar, and Tal Hassner. Age and gender estimation of unfiltered faces. IEEE Transactions on Information Forensics and Security, 9:2170–2179, 2014.
- [2] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks, 2018.
- [3] Fazle Rabbi. Facial age dataset. https://www.kaggle.com/datasets/frabbisw/facial-age.
- [4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.
- [5] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks, 2020.
- [6] K. Zhang, W. Liu, Z. Yu, and Z. Li. Age classification using deep convolutional neural networks. IEEE Transactions on Image Processing, 26(11):5587–5597, 2017.