

## Task 1

### Question 1

#### Intro

- The primary focus of this study is to analyse whether there is a way to improve a model's accuracy rates when an image in the dataset is distorted in some way. Convolution neural networks will be employed in this study, making use of pre-trained models such as VGG and Resnet. Transfer learning will be used to try to improve the accuracy of the models. The research will try to answer the following questions.:
- How can CNN technology be used to help improve accuracy in low resolution or distorted images?
- Can adding distortions to a dataset improve the accuracy of a given model?
- Does adding multiple distortions to a dataset help increase accuracy even further?

### Question 2

Attached Excel book

### Question 3

#### Literature Review

#### **Review of datasets used in other projects**

- During the literature research, several datasets were investigated, including Labeled Faces in the Wild (LFW), a public dataset, the VGGFace 2 Dataset, and the UCCS dataset, which is private.
- I noticed a distinction between private and public datasets. Some private datasets can be quite helpful for particular studies because they are case-specific and domain-specific. However, some researchers prefer public datasets since the study can be replicated. If they compare data to a study, they can verify the data if the dataset is

public; otherwise, gathering such data would be impossible (Hu et al., 2015.). When competitions like the International Challenge on Biometric Recognition in the Wild are held, public datasets are used so that the results from each participant can be reproduced. (ICB-RW)

### **Review of basic terminology**

- Basic terminology was researched so that a further understanding of the basics of the study are well understood. For instance, it was interesting to find out the relationship between Machine Learning Artificial Neural Networks and Convolutional neural networks.

### **Review of different approaches**

- As various researchers worked to improve the accuracy of facial recognition applications , the International Challenge on Biometric Recognition in the Wild (ICBRW) proved to be a valuable resource. Several ways to improve facial recognition accuracy have been attempted, and these approaches could all be studied. For example, Ghaleb et al. (2018) won the competition with an accuracy rate of 85.3 per cent in Rank 5 using VGG, an approach using CNNs. In contrast, Wang et al. attempted to boost accuracy by utilising super-resolution in the same year but only achieved a 59.04 per cent accuracy rate.
- I encountered different research approaches through the ICB-RW competition. Some researchers opt to create different models. For instance, Wu et al. (2017) developed the Pulse Coupled Neural Network (PCNN), a model based on the synchronised oscillating pulse of a cat's visual brain. Through the competition, I was able to find a substantial amount of research that worked on similar projects that helped me find studies related to my area of research.

## **Review of different metrics used in studies**

- The accuracy of face recognition in different types of datasets such as datasets from CCTV cameras used in studies (Aghdam et al., 2019) and datasets such as the LFW also frequently used in studies (Punnappurath et al., 2015) were analysed so that an appropriate dataset could be chosen.
- Distortion types of gaussian blur, box blur, and noise affect accuracy rating in facial recognition applications was studied by Karahan et al. (2016). These types of studies were studied to better understand what metrics need to be tracked.

## **Research Methodology**

### **Researched the CNN's that are going to be used**

- The VGG, ResNet50, and Xception were chosen after extensive research regarding various CNNs.
- The VGG16 and ResNet50 models were selected since they were frequently mentioned in the literature study and would be a suitable way to compare findings in the discussion of results(Aghdam *et al.*, 2019). The Xception model was chosen since it is the most recent version of Google's CNN and to see if it can considerably enhance results in this study and how it compares to studies in the literature review.

### **Prototype Research**

- After establishing the prototype structure, research was carried out into how to implement a CNN using code.

- During this research period, YouTube proved to be a valuable resource, as numerous videos demonstrate how CNNs are implemented. This research determines which libraries are the best to utilise based on this type of study and the distinctions between individual libraries, such as complexity and user-friendliness.
- Keras was chosen since it is a powerful library that lets you implement several CNNs with the same code. The Keras library was also picked because it has extensive documentation that is relatively easy to implement, as shown by the numerous examples provided in the documentation.

## Validity

- The validity and reliability of the data generated by the study were also considered. As Creswell (Research Design 4th edition, 2009) points out, this stage is critical in experimental research since certain safeguards must be in place to ensure that tests are fair. One of the methods used in this study was to train the CNNs many times and then average the results.

## References

Hu, G., Yang, Y., Yi, D., Kittler, J., Christmas, W., Li, S.Z. and Hospedales, T. (2015). *When Face Recognition Meets with Deep Learning: An Evaluation of Convolutional Neural Networks for Face Recognition*.

Wu, J., Yin, J., Zhang, Q. and Institute of Electrical and Electronics Engineers. (2017). *Proceedings of IEEE 13th International Conference on Electronic Measurement & Instruments* : Oct. 20-22, 2017, Yangzhou, China.

Proença, H., Nixon, M. and Nappi, M. (2018). Trends and Controversies. [online] [ieeexplore.ieee.org](https://ieeexplore.ieee.org). Available at: <<https://ieeexplore.ieee.org/document/8423530>> [Accessed 5 March 2022].

Karahan, amil, Kılınç Yıldırım, M., Kırtaç, K., Rende, ukrü, Bütün, G. and Kemal Ekenel, H. (2016.). *How Image Degradations Affect Deep CNN-Based Face Recognition?*

Aghdam, O.A., Bozorgtabar, B., Ekenel, H.K. and Thiran, J.P. (2019), "Exploring factors for improving low resolution face recognition", *IEEE Computer Society Conference on*

*Computer Vision and Pattern Recognition Workshops*, Vol. 2019-June, IEEE Computer Society, pp. 2363–2370.

Punnappurath, A., Rajagopalan, A.N., Taheri, S., Chellappa, R. and Seetharaman, G. (2015), "Face Recognition Across Non-Uniform Motion Blur, Illumination, and Pose", *IEEE Transactions on Image Processing*, Institute of Electrical and Electronics Engineers Inc., Vol. 24 No. 7, pp. 2067–2082.

#### Question 4

##### **Evaluation of Prototype**

- The effectiveness of the approach using the prototype. The results of when the CNN is trained using a dataset with distortions vs a model that has been trained on a dataset that doesn't contain any distortions. The accuracy will be considered for both approaches.
- The efficiency of the prototype. The amount of time it takes to train the dataset and produce results for one model while also commenting on the amount of time all of the tests took to complete.

##### **Interpreting Results**

- To assess the prototype's accuracy, the first test will be conducted with a dataset that does not contain any distortions, ensuring that the CNN's ground accuracy was accurate. So that it may be compared fairly against datasets that have been distorted. This will be done in order for the study's data to be as accurate as possible.
- The data from the prototype's outcomes will be displayed in diagrams. Each diagram will provide all necessary information, such as the type and amount of distortion added to the dataset. A short paragraph will be placed below so that readers may acquire all of the information they need from that paragraph rather than having to browse through the study to find the relevant diagrams.
- After analysing the results it is important to address whether the hypotheses or questions were supported or whether they were refuted as discussed in Creswell's book *Research Design* (4<sup>th</sup> edition 2019). A conclusion needs to be made on the reason why you think the research was successful or the reason why it was not successful.

- In a study by Aghdam et al. (2019), a similar approach was taken where the images in the gallery(dataset) were down-sampled to attempt to increase the accuracy in images with distortions. It will be a study that will be focused on in the discussion of results to see if an increase was also achieved and by how much the accuracy would have improved over models that were not trained with distortions.

## **Study Evaluation**

- Discuss if the approach taken in the study was successful or unsuccessful and the reason for both. Discuss if there is an approach that could be taken that is more efficient or if you would add anything to your approach to improve the results
- Could there be improvements to the dataset or maybe use a different dataset.
- After analysing the data are there any findings that were of interest that were not initially mentioned in the hypothesis.

## **Conclusion**

- The main findings of the study will be discussed, such findings might include the average accuracy by which a model was improved when trained with a dataset that included one or multiple distortions.
- Mention possible limitations such as not having enough computing power to increase certain parameters for models.
- The shortcomings of the study will be discussed.
- Any future recommendations for the study. One possible recommendation would be to create a model and adjust all the weight to fit your use case rather than using transfer learning.

## **Introduction**

- Research into the basics of face recognition
- Mention the main areas of research that will be going to be studied
- Research information on face recognition in images that are not ideal.

## **Abstract**

- Introduction of the subject being studied
- The main findings discovered in the study

## **References**

Aghdam, O.A., Bozorgtabar, B., Ekenel, H.K. and Thiran, J.P. (2019), "Exploring factors for improving low resolution face recognition", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Vol. 2019-June, IEEE Computer Society, pp. 2363–2370.

## Question 5

## Conclusion

The Literature review and the methodology of the dissertation are completed however after working further with the prototype these might need to be altered to include other information which I was not previously aware of. The discussion of results will be the next phase of this project it will go into the results after adding the distortions to the dataset. I will try to keep up with the deadlines stipulated in the Gannt chart however in the previous semester it was not possible to keep with the stipulated timeline during the peak examination months.

## Task 2

[https://mcastedumy.sharepoint.com/:u:/g/personal/matthew\\_grima\\_c10283\\_mcast\\_edu\\_mt/EQAo6Xn70ypCoMs3S3xoG04Be0gsAxYEzuXaEhm3PJhDyw](https://mcastedumy.sharepoint.com/:u:/g/personal/matthew_grima_c10283_mcast_edu_mt/EQAo6Xn70ypCoMs3S3xoG04Be0gsAxYEzuXaEhm3PJhDyw)

## Task 3

1. Quantitative study The purpose of this **experimental** study **will be** to test the theory of **face recognition** that **relates to** the **low-resolution images**(independent variable) to the **CNN model's face recognition accuracy rate**(dependent variable). The independent variable(s) **low resolution** will be defined as **images taken under mismatched conditions meaning images taken under conditions where the face is at an extreme pose, occlusions are present such as sunglasses, various illumination, motion blur or focus** (provide a definition). The datasets used the VGGFace2 and the MS-Celeb-1M Dataset are images taken in the wild meaning they were taken under a variety of different conditions. The dependent variable(s) will be defined as the **accuracy of the models** (provide a definition)

Aghdam, O.A., Bozorgtabar, B., Ekenel, H.K. and Thiran, J.P. (2019), "Exploring factors for improving low resolution face recognition", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Vol. 2019-June, IEEE Computer Society, pp. 2363–2370.

2. Quantitative study The purpose of this **experimental** study **will be** to test the theory of **face recognition** that **relates to** the **degraded images**(independent variable) to **CNN model's face recognition accuracy rate**(dependent variable), the **controlling factor in this study was the amount of distortion added to the images**. The dataset set used

called the labelled faces in the wild is a dataset that has images of individuals in the wild meaning they are taken under a variety of different environments. The independent variable **degraded images** will be defined as **images that have distortions added to them such as noise and blur** (provide a definition). The dependent variable will be defined as the **accuracy of the models**

Karahan, amil, Kılınç Yıldırım, M., Kırtaç, K., Rende, ukrü, Bütün, G. and Kemal Ekenel, H. (2016.). *How Image Degradations Affect Deep CNN-Based Face Recognition?*

3. Quantitative study The purpose of this **experimental** study will be to test the theory of **face recognition** that **relates to** (describes outcomes) or (compares? relates?) the **images taken from surveillance cameras** (independent variable) to the **accuracy of the VGG CNN model**(dependent variable). The dataset which is called the ICB-RW is made up of images taken from CCTV cameras. The independent variable(s) **images with extreme pose** will be **defined as images that have been taken from CCTV cameras** (provide a definition). The dependent variable(s) will be defined as the **accuracy of recognising individuals using the VGG CNN model** (provide a definition)

Ghaleb, E., Ozbulak, G., Ekenel, H. k and Gao, H. (2018), "DEEP REPRESENTATION AND SCORE NORMALIZATION FOR FACE RECOGNITION UNDER MISMATCHED CONDITIONS", *IEEE Intelligent Systems*, Vol. 33, pp. 43–46.

#### Task 4

##### Research Design

The area of research being investigated is facial detection in images that have been distorted by noise or blur. The study's key objective will be to improve the accuracy with which a face can be identified in a distorted image using Convolutional Neural Networks. Throughout the investigation, the accuracy of the model trained with CNN technology will be a significant factor. Because of the aforementioned points, this study will be quantitative, as its sole goal is concerned with the accuracy of the CNN models, and this cannot be measured using other forms of research design such as qualitative or mixed research. Data gathering, analysis, interpretation, and report writing using qualitative methods differ from typical quantitative methods. Qualitative methods include deliberate sampling, open-ended data gathering, text or image analysis, information representation in figures and tables, and personal interpretation of the findings. In a mixed-methods study, quantitative and qualitative data are collected and "mixed" or integrated. It might be understood that both sorts of data provide different data types (open-ended data in qualitative and closed-ended data in quantitative). The capabilities of each type of data collection can be combined to generate a better knowledge of the topic



being studied if we assume that every kind of data collection has both limitations and strengths (and, as well, overcome the limitations of each). It might be claimed that this "mixing" of facts provides a better comprehension of the problem or question than either method used alone. However, since the dissertation is mainly concerned with the accuracy rate of a convolution neural net it is not possible to have a mixed approach. All of the studies cited in the dissertation used a quantitative research design, which is the only technique to conduct such a study. For instance, Aghdam et al. (2019) Attempted to increase face recognition accuracy in models by adding distortions to images with which the model was trained, the researchers made use of transfer learning. In this study, the research design chosen to gather data was also quantitative since they wanted to measure the accuracy rates of the model. A study conducted by Yildirim et al (2016) studied the effect distortions have on different types of models. The same images were used on different types of models and the accuracy rates of all the models were measured to see which models are affected the most. This study also used a quantitative research design so that it can measure the accuracy rates of the model and other important data such as the performance i.e how long the models took to analyse the different types of distortions. Face recognition using CNN technology is well studied, however, there has been far less research on using CNN technology to identify people in images that were not captured under ideal conditions. Only a few studies have adopted an approach comparable to the one that will be employed in this study. This study will try to increase the accuracy of identifying individuals in distorted images by adding one or more distortions to the image of the subject inside the dataset. This approach is a justification for the research design chosen as there is a lack of studies using this approach as explained above. Several different models will be created for this study which can be considered as experiments. As discussed by Cresswell (2014) true experiments that take into consideration independent and dependent variables both of which are being considered in this study should be studied using the Quantitative research design.

## **Research Methods**

There are various approaches to solving the task of recognising individuals in distorted images, it was decided to focus on CNN's because it appeared to be the dominant technology in the studies researched. Such an example is, the winning team in the 2016 Challenge on Biometric Recognition in the Wild (in the wild implies photographs collected in less-than-ideal conditions) used CNNs (Ghaleb et al , 2016). Other contestants attempted to apply technology such as a Simple Neural Net (SLM) in the competition, but their accuracy rates were significantly lower (Brogan, 2016). After the competition, a number of methods were tried to enhance the accuracy rate, for example, one researcher employed super-resolution to improve the image quality of the distorted images(Wang et al. 2016). While another researcher created a model based on the phenomena of a cat's visual cortex's synchronised oscillating pulse. (Feifei et al. , 2017). Later experiments that were successful in greatly improving the accuracy rate from 2016 did so by using more complicated CNNs and fine-tuning their model (Aghdam et al. , 2019). Another study that took place before the competition in 2015 sought to develop a model that was particular to different sorts of distortions this was very effective(Punnappurath et al.); however, this technique would require multiple different models for each distortion, which would be impractical. As a result, the approach used in this study aims to be independent of

image distortion, making it a more versatile approach. Other technologies such as SVM were also used and had significant results at completing similar tasks being tackled in this study however during the research period of the dissertation there weren't enough studies found using this technology to consider using it.

The dataset for this study will consist of over 1000 images, which is significant for conducting this type of research as each individual must be assigned to a separate category. Because a similar number of images were used in the ICBRW competition, this dataset would be considered sufficient. However, the model must be fine-tuned before it can be trained and tested using this dataset. Because a similar number of photographs were used in the ICBRW competition, this dataset would be considered sufficient. However, the model must be fine-tuned before it can be trained and tested using this dataset. Fine-tuning includes deleting images with occlusions like sunglasses, images with faces at extreme angles, and images with more than one face in them. All of these images were removed because the study's major focus is on distortions such as noise and blur, rather than the other external influences described earlier.

Each model will be trained and evaluated with images from the dataset that include no distortions in order to obtain an accuracy rate for the model chosen. This process will be executed five times on the same dataset in order to calculate an average. This is done so that the accuracy of the dataset chosen without distortions may be compared to the datasets with distortions, and any discrepancies in accuracy can be identified.

## Research Pipeline

All of the pipeline's phases were separated and placed at the top of the design so that they could be clearly viewed and understood. The project's primary elements were placed on the left side of the diagram's design and were each given colour. The arrows can also easily be followed with this layout as they are mostly in straight lines which is another advantage of this layout. The layout of this pipeline design should make it simpler for readers to comprehend the pipeline.

The process will begin with the dataset being divided into two folders. The identical image of an individual cannot be found in both the test data and the training folders. Various amounts of combinations and degrees of distortions will be added to both the test and train images, the paths will be set to the training and testing images folders. The CNN chosen will be implemented with all the trainable layers set to false since transfer learning will be used in this study. The CNN model's settings will be determined in terms of cost, optimisation, and the metrics that will be tracked, which in this case will be the model's accuracy. Image data will be created, which means that the same images in the folders will be copied, scaled, and flipped to provide more images for the model to analyse. The correct batch number and epochs number,

which will primarily affect the test accuracy, will guarantee that the model is not under or overfitted. The data for each epoch's will be shown however a chart will be shown that shows the progress of the test and train data from the start to finish.

## References

Aghdam, O.A., Bozorgtabar, B., Ekenel, H.K. and Thiran, J.P. (2019), "Exploring factors for improving low resolution face recognition", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Vol. 2019-June, IEEE Computer Society, pp. 2363–2370.

Feifei, S. , H. ,M, Rencan, N., Wu , Z. Institute of Electrical and Electronics Engineers. (n.d.). Proceedings of IEEE 13th International Conference on Electronic Measurement & Instruments : Oct. 20-22, 2017, Yangzhou, China.

Ghaleb, E., Ozbulak, G., Ekenel, H. k and Gao, H. (2018), "DEEP REPRESENTATION AND SCORE NORMALIZATION FOR FACE RECOGNITION UNDER MISMATCHED CONDITIONS", *IEEE Intelligent Systems*, Vol. 33, pp. 43–46.

Wang, Z., Chang, S., Yang, Y., Liu, D. and Huang, T.S. (2016). *Studying Very Low Resolution Recognition Using Deep Networks*.

Punnappurath, A., Rajagopalan, A.N., Taheri, S., Chellappa, R. and Seetharaman, G. (2015), "Face Recognition Across Non-Uniform Motion Blur, Illumination, and Pose", *IEEE Transactions on Image Processing*, Institute of Electrical and Electronics Engineers Inc., Vol. 24 No. 7, pp. 2067–2082.

Karahan, amil, Kılınc Yıldırım, M., Kırtac, K., Rende, ukrü, Bütün, G. and Kemal Ekenel, H. (2016.). *How Image Degradations Affect Deep CNN-Based Face Recognition?*