Get more from less: Differential neural decoding for effective reconstruction from limited training data

Thirza Dado & Umut Güçlü

Neural Coding Lab, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands

DONDERS INSTITUTE

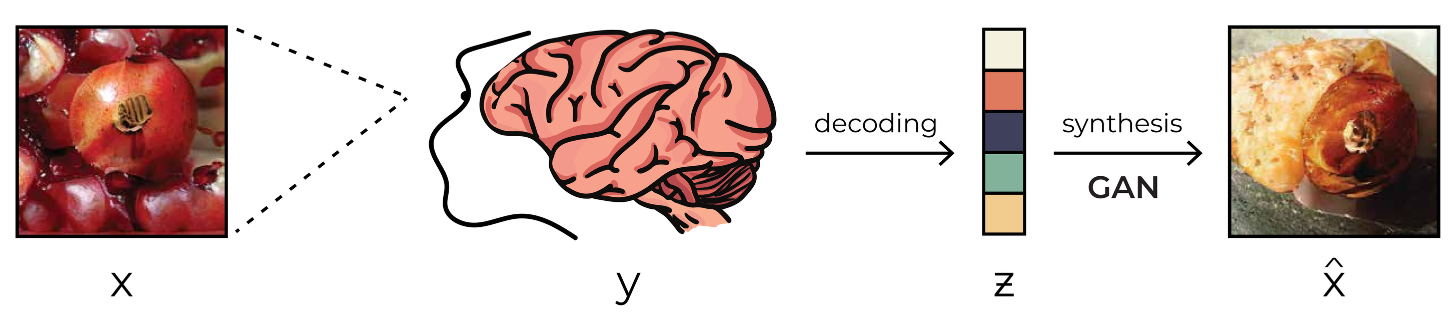
INTRODUCTION

• Neural decoding refers to the computational problem of classifying or reconstructing perceptual stimuli, $\mathbf{x}$, based on stimulus-evoked neural activity, $\mathbf{y}$.

Problem:
• Biological sensory processing consists of highly nonlinear transformations
• Modeling requires extensive datasets of stimulus and neural response pairs
• Neural responses exhibit trial-to-trial variability and noise (to identical stimuli)

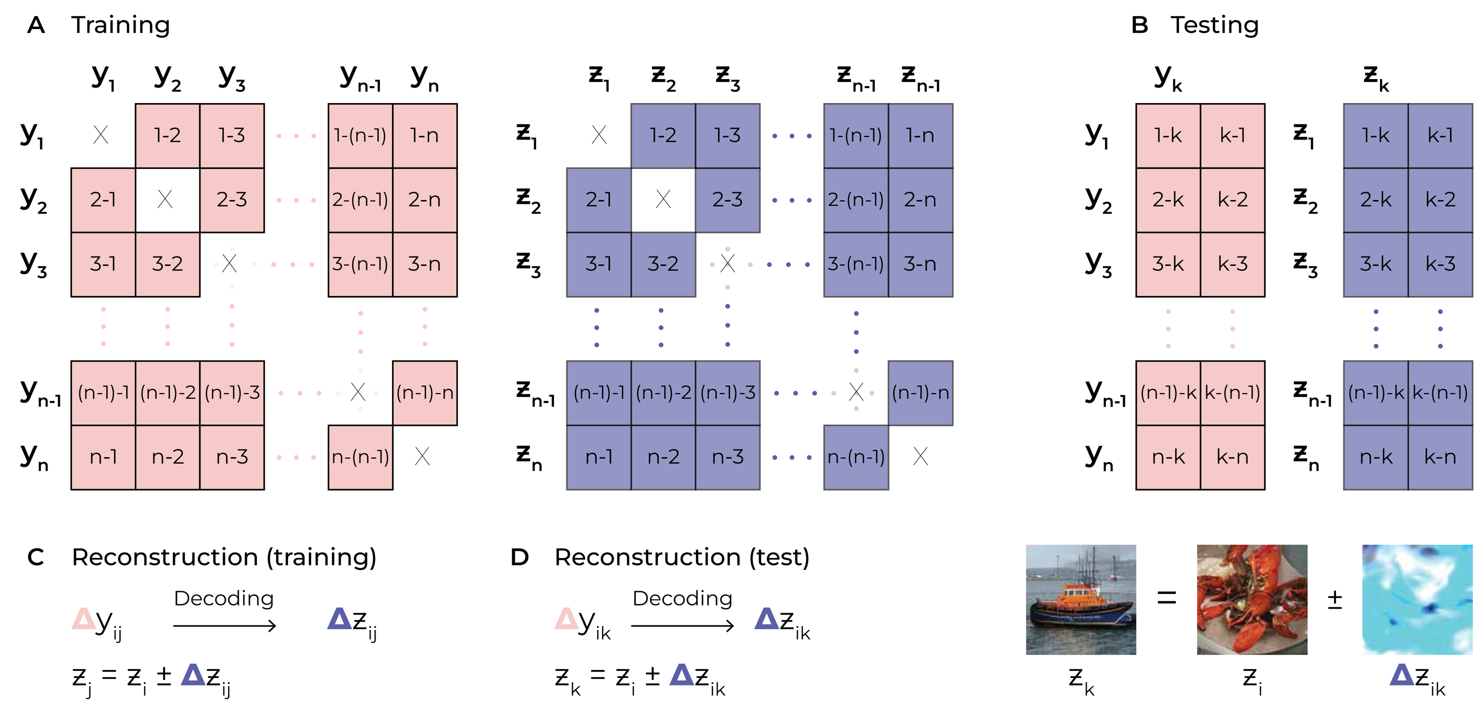
Solution:
• Generative adversarial networks (GANs) can synthesize photorealistic images from latent codes, $\mathbf{z}$, [1] which can be used in neural decoding as intermediate feature representation between neural activity and reconstruction, $\hat{\mathbf{x}}$ [2, 3]

• We propose a new framework for DIfferential neural deCODING (**dicoding**) that reconstructs stimuli from *relative differences* between pairs of neural responses instead of directly predicting single latents from single responses



METHODS

Given a dataset [3] of neural responses, $\mathbf{y}$, to corresponding stimuli, $\mathbf{x}$, and their underlying latents, $\mathbf{z}$, we compute all pairwise differences, $\Delta\mathbf{y}_{ij} = \mathbf{y}_i - \mathbf{y}_j$ and $\Delta\mathbf{y}_{ij} = \mathbf{y}_j - \mathbf{y}_i$, resulting in a quadratic expansion of the training dataset. We then train a decoder model to map these response differences to latent differences, $\Delta\mathbf{z}_{ij}$.

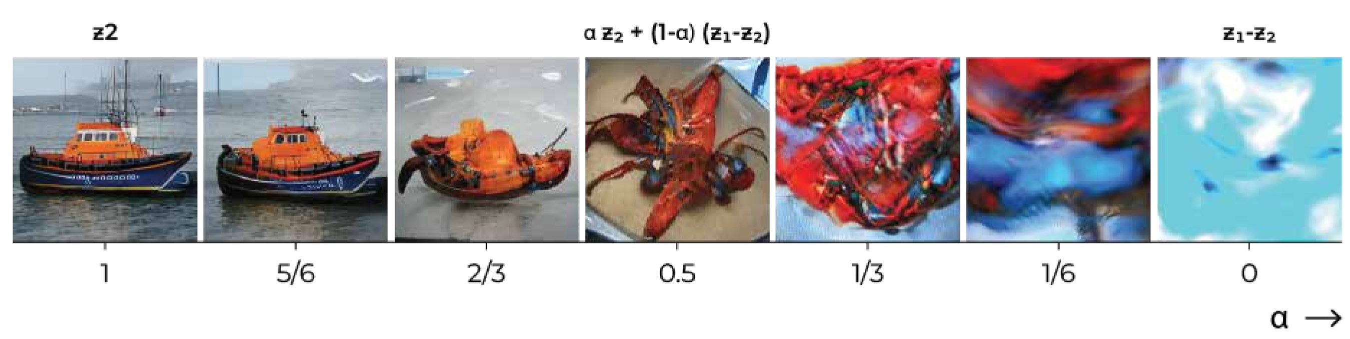


THEORETICAL ANALYSIS

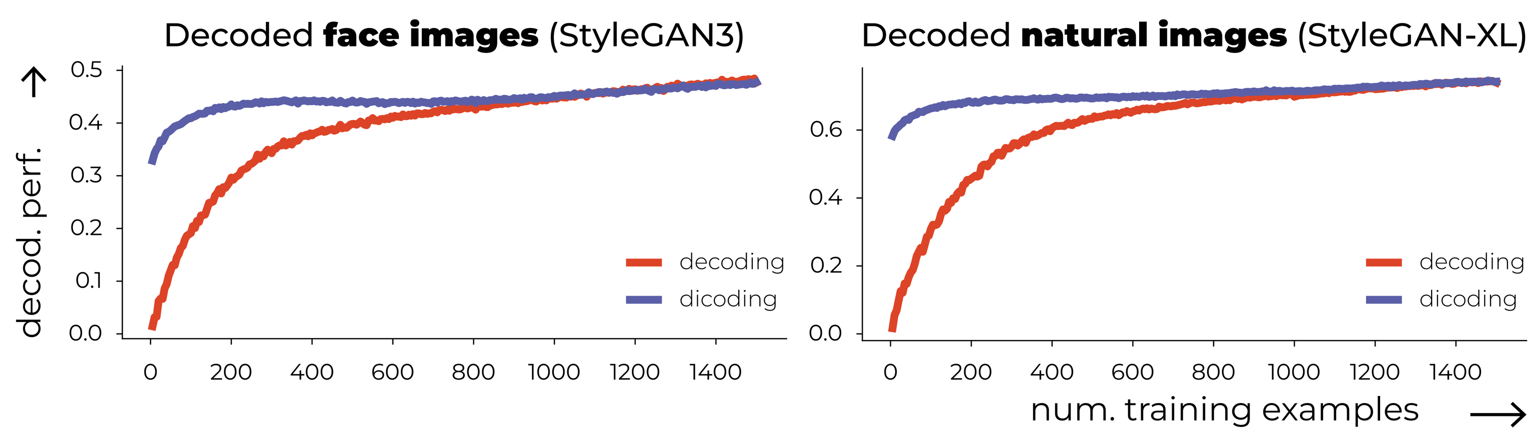
• **Noise robustness.** We model noise as independent additive Gaussian with covariance $\Sigma = \sigma^2\mathbf{I}$: $\mathbf{y}_i = f(\mathbf{z}_i) + \varepsilon_i$, $\mathbf{y}_j = f(\mathbf{z}_j) + \varepsilon_j$, where εi, εj in $\mathcal{N}(0, \sigma^2\mathbf{I})$ are independent noise. The difference $\delta\mathbf{y}_{ij} = \mathbf{y}_i - \mathbf{y}_j$ retains this additive structure: $\delta\mathbf{y}_{ij} = \delta\mathbf{z}_{ij} + \delta\varepsilon_{ij}$ where $\delta\varepsilon_{ij} = \varepsilon_i - \varepsilon_j$ has twice the variance $2\sigma^2\mathbf{I}$ of the individual terms. That is, as the data increases quadratically, the uncorrelated noise increases linearly.

• **Latent space exploitation.** We assume that the GAN generator's latent space is an isotropic Euclidean metric space with orthogonal disentangled axes spanning independent perceptual factors which enables seamless geometric manipulations. In this idealized setting, latent difference vectors $\delta\mathbf{z}_{ij} = \mathbf{z}_i - \mathbf{z}_j$ precisely capture stimulus differences $\delta\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ without interactions.

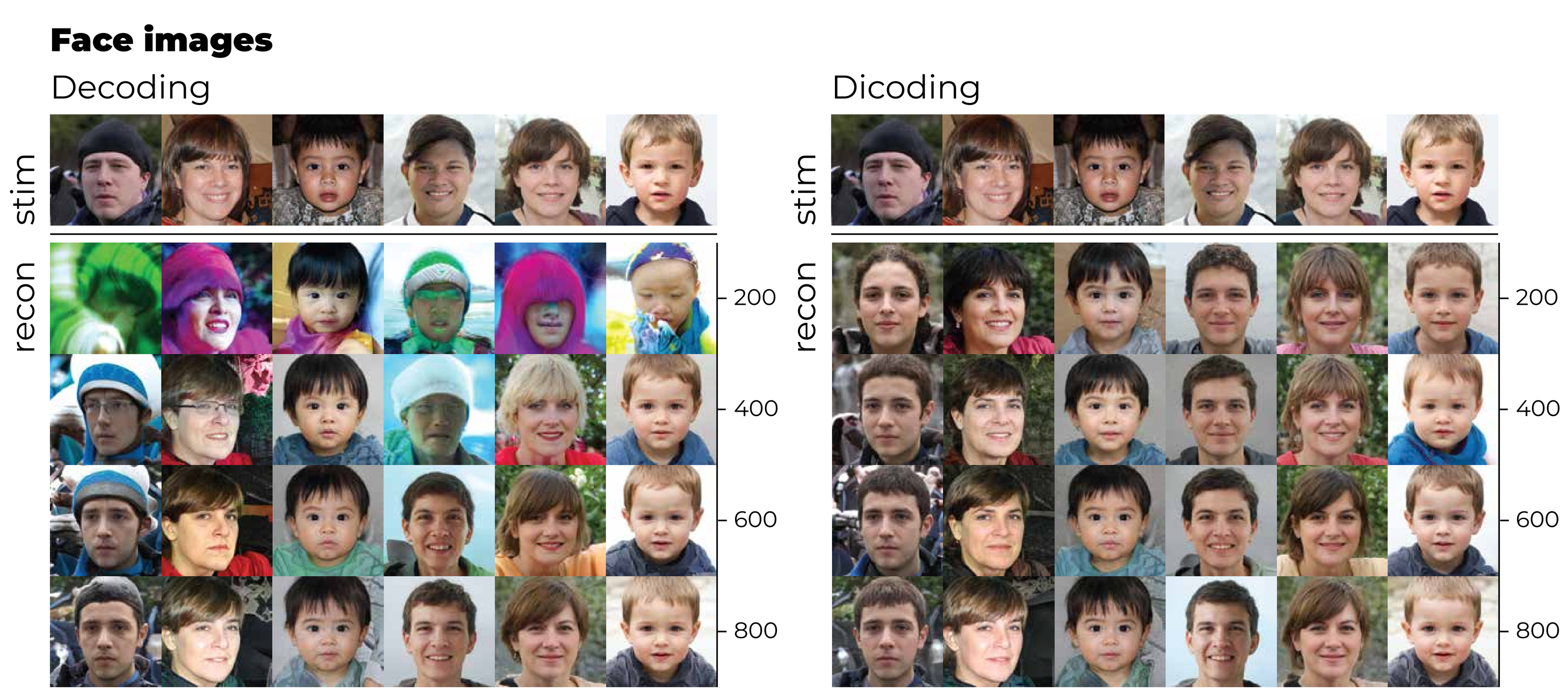
• **Improved sample efficiency.** We assume that the latent space exhibits proportionality; the magnitude of latent difference vectors $\delta\mathbf{z}_{ij}$ scales linearly with stimulus semantics $\delta\mathbf{x}_{ij}$. Crucially, this implies that doubling the training set from n to $2n$ not only expands the size of the training set but also augments the volume spanned by the latent differences, providing broader coverage of potential variations within the latent space. The consequence is a more robust and effective training process as the model learns from a more diverse set of examples.

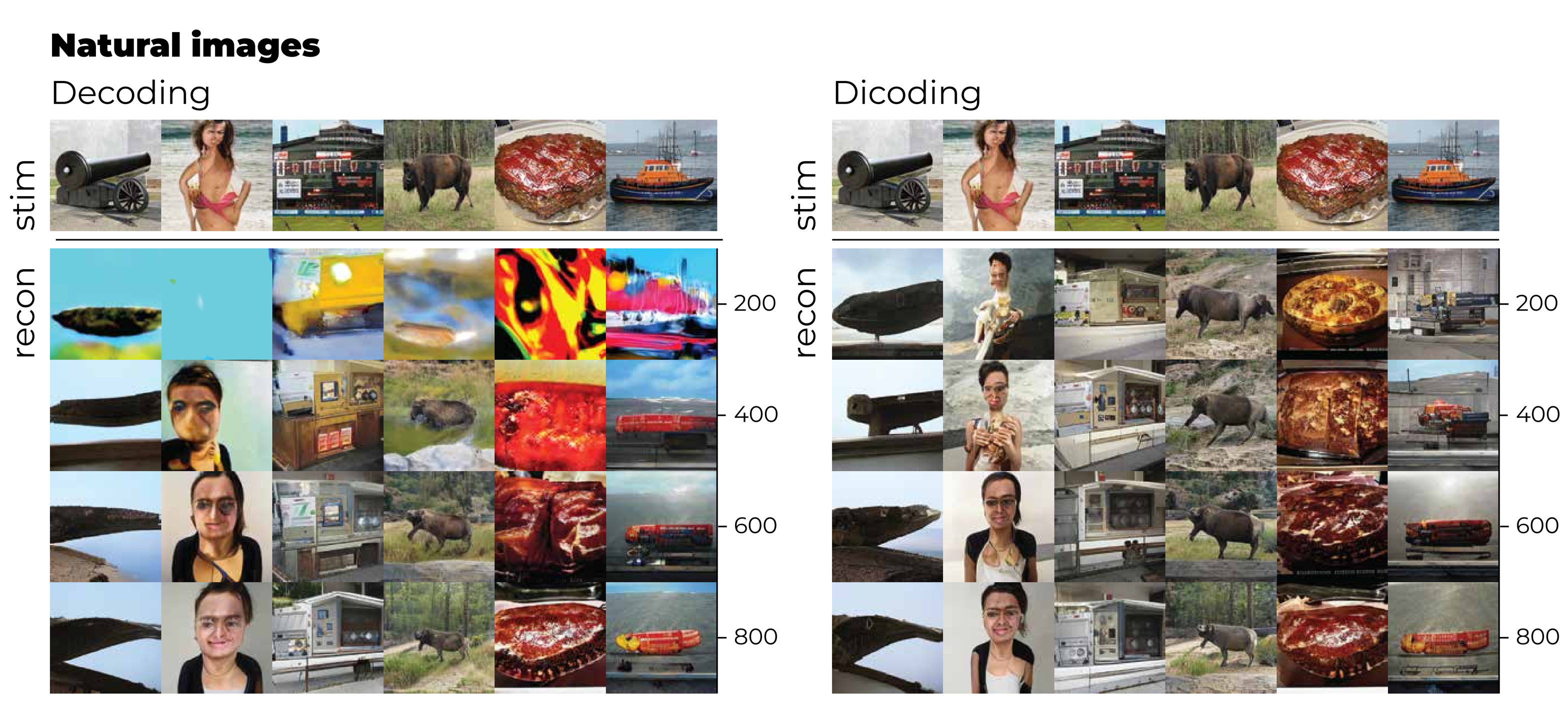


RESULTS

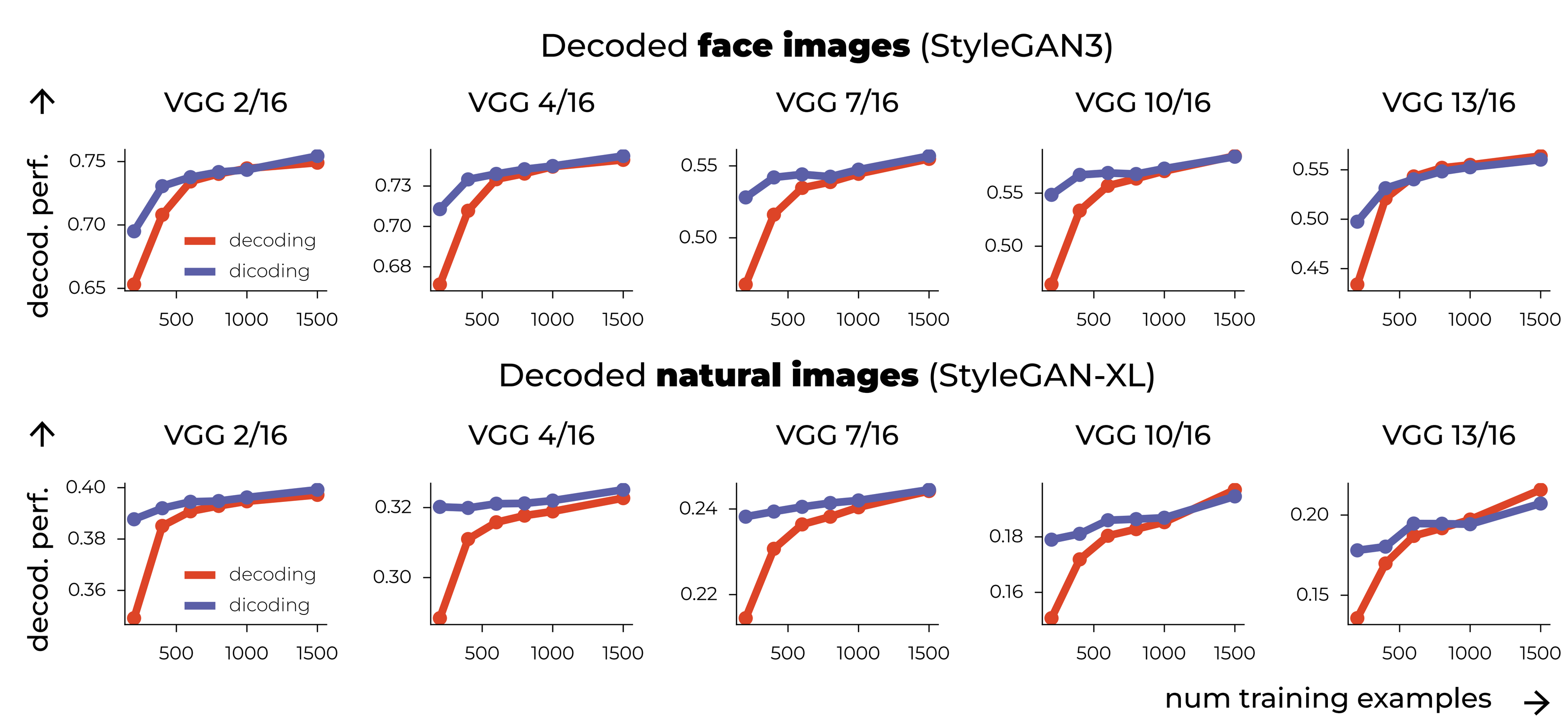


Decoding performance, in terms of cosine similarity between predicted and target latents, as a function of number of training examples. ***Dicoding*** *outperformed regular decoding for low samples sizes. Their performances converged as the training dataset size increased (~800).*





The predicted latents by the decoders trained on 200, 400, 600, and 800 training faces and images were fed to their respective GAN generator for the reconstruction of the corresponding images. Notably, ***dicoding*** *reconstructed higher-quality images for limited sample sizes.*



Perceptual similarity using five activations of VGG16 for face recognition [4] in the case of faces and VGG16 for object recognition [5] in the case of natural images affirmed the superior performance of ***dicoding*** *for low sample sizes.*

CONCLUSIONS

Our theoretical analysis demonstrated potential benefits in noise robustness, latent space exploitation and data augmentation after which empirical analysis validated these advantages in practice for low amounts of training data.

As such, our **dicoding** paradigm:

• successfully mapped response differences to latent offsets which effectively encoded meaningful directions in latent space, leveraging the inherent semantic geometry thereof

• is poised to help enable the next generation of neural interfaces and prosthetic devices through more effective utilization of limited biological data and meaningful semantic transformations grounded in statistics of real-world data

REFERENCES

[1] Goodfellow, P. A., Mirza, X., & Warde-Farley, O. (2014). Goodfellow I. Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., Generative adversarial nets, *Advances in neural information processing systems, 27.*
[2] Dado, T., Güçlütürk, Y., Ambrogioni, L., Ras, G., Bosch, S., van Gerven, M., & Güçlü, U. (2022). Hyperrealistic neural decoding for reconstructing faces from fMRI activations via the GAN latent space. *Scientific reports, 12(1), 141.*
[3] Dado, T., Papale, P., Lozano, A., Le, L., Wang, F., van Gerven, M., ... & Güçlü, U. (2023). Brain2GAN: Feature-disentangled neural coding of visual perception in the primate brain. *bioRxiv, 2023-04.*
[4] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.*
[5] Parkhi, O., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. In *BMVC 2015-Proceedings of the British Machine Vision Conference 2015.* British Machine Vision Association.

Thirza Dado; thirza.dado@donders.ru.nl; https://neuralcoding.nl/

Radboud University Radboudumc