Enhanced Communication for ALS Patients

A Technical Report submitted to the Department of Biomedical Engineering

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Deyan Saleem

Spring, 2025

Technical Project Team Members

Kunal Bahl

Ishaan Shah

Ali Nilforoush

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Timothy E. Allen, Department of Biomedical Engineering

Alec Bateman, Barron Associates, Inc.

Michael DeVore, Barron Associates, Inc.

Enhanced Communication for ALS Patients

Kunal D. Bahl[a], Ishaan Shah[a], Deyan Saleem[a], Ali R. Nilforoush[a]

[a] Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22904

Abstract

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative disease causing a loss of motor function, including breathing and speaking impairment. Late-stage ALS patients lose the ability to communicate entirely. To aid in speaking, Augmentative/Alternative Communication (AAC) devices have been developed that use eye gaze and gesture devices to form signals. However, while AACs that rely on eye tracking provide a valuable communication tool for ALS patients, the use of BiPAP masks for breathing support often obstructs the eyes, hindering their ability to interact with these systems effectively. In addition to visual occlusion, current tablet-based AAC systems are often impractical for continuous use, especially when the patient is sleeping or lying down, limiting their reliability in spontaneous communication scenarios. A blink-detecting camera attachment was designed to BiPAP masks to enable continued communication despite obstructions. The device integrated a wired camera, camera mount for BiPAP masks, Arduino board, and eye-state detection algorithm. For the algorithm, a convolutional neural network (CNN) was utilized and trained to classify between open and closed eyes on open-source eye image data (>95% validation accuracy). The CNN detects eye-states real-time and triggers a virtual button press (VBP) when a closed eye state is detected >2 seconds. Preset combinations of VBPs will be programmed to generate communication signals (i.e., “Yes,” “No,” “Hello”). Prototype is currently being finalized, and testing on healthy and ALS subjects will be conducted in the future.

Keywords: **CNN, Computer Vision, Blink Detection, CAD, AAC, ALS**

Introduction

What would you do if you were in the hospital with an emergency and physically couldn’t ask for help? How would that make you feel? Aside from emergency scenarios, the ability to communicate is paramount to the livelihood of every person and has been shown to directly correlate with a patients quality of life (QoL) (Felgoise et al., 2016). Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease involving motor neuron degeneration resulting in muscle weakness, respiratory failure, and loss of the ability to communicate. While the cause is exactly unknown, current possible mechanisms include genetic mutations resulting in an aggregation of the SOD1 protein, oxidative stress, glutamate excitotoxicity, and hereditary involvement (Brotman et al., 2024). The respiratory weakness of patients inhibits proper breathing and contributes to hypercapnia, which can result in sleep disturbance, fatigue, and depression. Non-invasive ventilation (NIV) was developed to treat this issue by providing ventilatory support through the use of a bilevel positive airway pressure (BiPAP) mask (Dorst & Ludolph, 2019).

Augmentative Alternative Communication (AAC) devices have been developed to enable ALS patients to communicate with their physicians and caregivers. For example, some AAC devices translate eye tracking into letters and numbers on a control board (Ezzat et al., 2023). AAC systems that rely on eye tracking provide a valuable communication tool for ALS patients; however, the use of BiPAP masks often obstructs the eyes, hindering their ability to interact with these systems effectively.

Current AACs focused around blink/eye tracking have a strong backbone within the literature in terms of algorithm development. Work done by Dewi et al. (2022) classifies blinks based upon a novel modified eye aspect ratio (EAR) and is able to achieve accuracies between 81% to 98% depending on the dataset and EAR used. Park et al. used a stacked hourglass convolutional neural network (CNN) to learn landmarks for gaze estimation, achieving among the lowest error rates at the time (2018). In total, the use of computer vision to classify the position or state of a human eye in real time is yielding great success and continuing to grow.

Developments in this field are important not only for restoring the communication to patients with ALS or who are otherwise unable to communicate, but also for the caregivers of the patient. Informal caregivers, untrained people taking care of a person with ALS (family member or friend), have been shown to experience a greater burden as ALS progresses (de Wit et al., 2018). This burden is in large part due to physical changes and depressive moods that may become onset throughout ALS.

However, a limitation that exists for many of the current eye-based AACs is that the patient must be looking directly into them. This makes sense when the patient is actively looking to communicate; however, if there is a situation in which the patient is unexpectedly met with a need for communication, current AACs don't have the omniscience to react to all situations. This concern is further compounded when the patient is wearing a BiPAP mask due to eye occlusion and camera field of view (FOV) impedance.

Current technology does not adequately address this limitation, particularly during nighttime use, leading to a loss of communication capabilities and an increased burden on caregivers. This product aims to integrate open-source blink detection algorithms, with an Arduino camera, and 3D-printed mask mounts with the end goal of converting a sequence of blinks into a form of communication. The combination of these technologies will allow for blink detection as a form of communication to be available to the patient wearing the BiPAP mask for as long as they have the mask on. Instead of the patient having to be presented with their AAC, it will always be readily available and positioned to a working angle. Therefore, in events of emergencies or times in which the need for communication cannot be predicted, the patient will still have a method for communication. As a secondary effect to having this AAC be readily available, caregiver burden is also partially mitigated. Caregivers will be able to grant patients with ALS privacy with the knowledge that should they need help, they have a consistent manner of asking for it.

Results

Product Aided Communication

The primary goal of this product is to serve as a primary communication tool for patients with ALS to use when they are unable to or inconvenienced to use their traditional AACs. The overall structure of the algorithm is depicted in Figure 1 and demonstrates a sample sequence of virtual button presses (VBPs) to signal 'I need help!' to a caregiver. A VBP is defined as a blink that takes at least 2 seconds. That is to say a period of at least two seconds in which the eyes are closed.

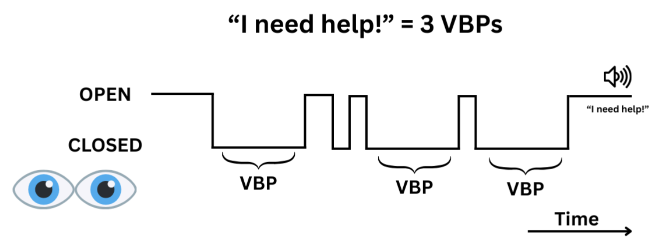

Figure 1. A high-level illustration of the algorithm's mechanism. The algorithm is a binary classifier and will either determine the eye as open or closed. When the eye is closed the algorithm will keep track of how long the eye has been closed for. If the eye is closed for longer than 2 seconds, that will count as a single virtual button press (VBP). The eye must be reopened before the next VBP is recorded. Different sequences of VBPs will indicate different predetermined phrases.

This timing was chosen because the activation of a VBP is intended to be intentional. If the system were to simply be based off of blinks then simply going about your day to day life would set off many undesired signals, and would likely cause the caregiver to be unable to distinguish between undesired signals and intended ones. Therefore the VBP, which requires a more intentional blinking pattern, is more resilient to this sort of mis-signal. Furthermore, a set of VBPs can be strung together to form a predetermined phrase. Two consecutive VBPs would bring up a YES or NO prompt to ask if a signal truly is intended to be sent, for which 1 VBP would indicate NO and 0 VBPs would indicate YES within a 10 second time period. Figure 1 demonstrates that a sequence of three consecutive VBPs would send a predetermined signal of 'I need help!' (which would also be confirmed by a similar YES or NO screen). The decision to afford only predetermined signals is reminiscent of a standard call light within a hospital. The call light is not itself a method of comprehensive communication, but rather a pre-communicative tool or an indication of a desire to communicate. Likewise this AAC is meant to simply communicate an emergency or sudden message that can then be handled more complexly by a caregiver or a more advanced AAC.

Blink Detection

The blink detection algorithm developed within this project drew inspiration from the software architecture of Dewi et al., 2022 and Park et al., 2018. The predominant structure is a loop between VBPs that will manually terminate (i.e. the patient chooses the NO prompt). Once the ArduCam starts recording the eye region, it will attempt to classify the eye as either open or closed, as informed by the CNN model. Once it is able to detect whether the eye is closed or open, it will continuously look for any moment in which the eye is closed for >2 seconds. This will then count as a VBP and once the eye is reopened, the cycle will repeat (Fig. 2). In this manner, the eye is being constantly monitored for a VBP.

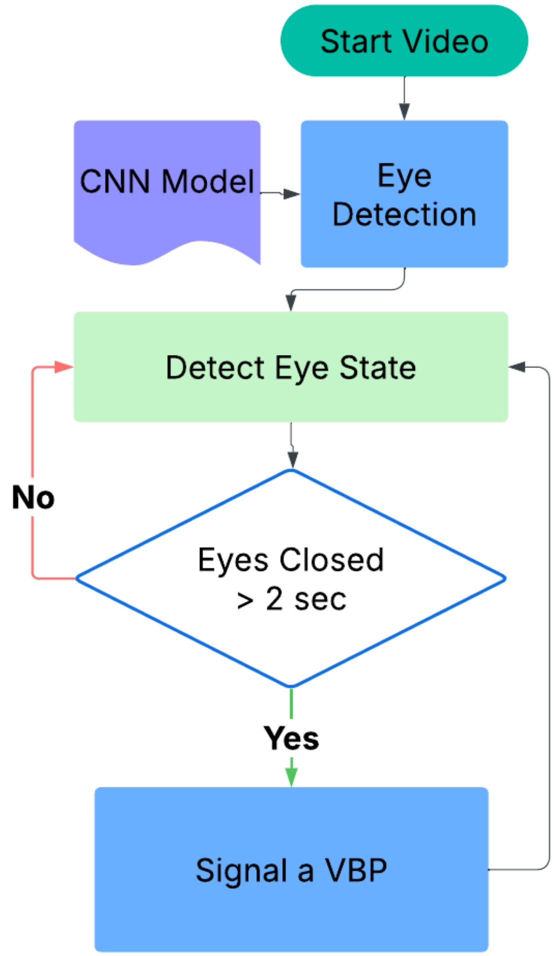

Figure 2. Flowchart overview of the blink detection algorithm used.

CNN

The 2-dimensional convolutional neural network (CNN) trained to classify whether the eye was open or closed utilized a three layer network and an input of 128 x 128 pixel black and white images. These images were sourced from the Media Research Lab and featured open and closed pictures of both the left and right eye (Fig. 3) (Jahan et al., 2023). In total, the initial training and validation dataset consisted of 2,372 images of each of the four types, 9,488 total images. From this data, a training and validation split of 90/10 (90% training and 10% validation) was used.

Each layer consisted of three main processes: filtering, max pooling, and batch normalization. The TensorFlow Keras module library was used for the construction of the CNN and each of the three layers had an increasing filter count– 32, 64, and 128 – filters representing the first, second, and third layers. Max pooling was conducted to downsample the spatial data for each of the inputted images based upon the maximum value within a sector of the 128 x 128 image space. Batch normalization finally normalized the output space.

However, the predominant task required from this CNN is the classification of whether the eye is open or closed when the patient is wearing a BiPAP mask. In lieu of a dataset collected via an IRB, a method of transfer learning from which the previous 9,488 images were used as a basic dataset was implemented. A dataset of 94 additional images with a more representative field of view (FOV) were then used to fine tune the previously trained model to utilize what the model already knew and further expand it to a FOV when the patient was wearing a BiPAP mask. This was done in a separate file to update the weights of the neural network with a much smaller learning rate. (Fig. 4).

The model was trained and validated over the course of 50 epochs and eventually achieved a validation accuracy rate of roughly 96% (Fig. 5). The validation accuracy curve (seen in red) experiences early fluctuations in accuracy which eventually peeters out in magnitude. This behavior is likely due to model overfitting to the training data set as well as the validation data not being representative of the whole dataset, which incurs errors that must be mitigated through updating weights. Consequently, the model was able to predict with between 90% to 96% accuracy after 50 epochs which is approximately where the desired outcome would be.

Figure 3. depicts sample images from Jahan et al. of open right (**a**), open left (**b**), closed right (**c**), and closed left eyes (**d**) used to train and validate the CNN (2023).

Figure 4. depicts a sample of the new data added post initial model training. Left image is of an open right eye. Right image is of a closed left eye.

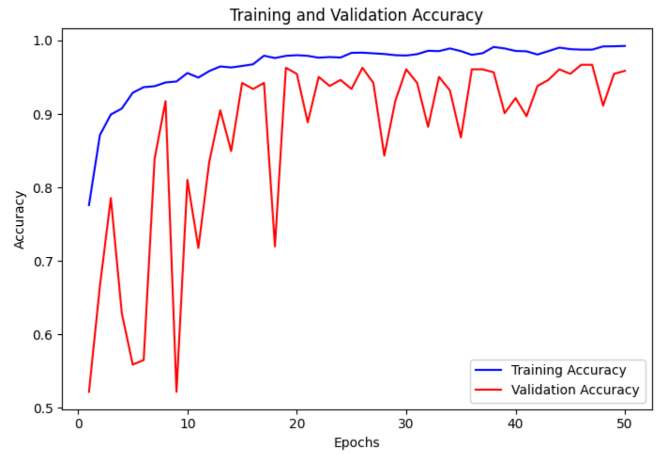

Figure 5. graphs depicting the training and validation accuracy of the 3 layer, 2-dimensional CNN model over the course of 50 epochs.

ArduCam Software

The functionality of the device is predicated upon the delivery of image data of the eye region to a processing unit. This task was achieved through the use of an ArduCam Mini 2MP Plus SPI camera module. The ArduCAM was programmed such that it was able to take photos at a rate of approximately 8 frames per second (fps). Then through the serial peripheral interface (SPI), the image data is transferred into a Python program that runs the calculations and actively attempts to predict whether or not the eye is open or closed (Fig. 6).

Figure 6. The blink detection algorithm classifies an eye as ‘Open’ (left image) and ‘Closed’ (right image) depending on the state of the eye. The image on the right was taken just as the eye was closed, so the proper > 2 second window needed to increment the VBP Count was not in effect yet.

Physical Apparatus

The physical apparatus for this project consists of a custom-designed camera mounting system integrated onto a BiPAP mask. This mount was specifically designed to facilitate precise eye-tracking for ALS patients who require respiratory support. To accomplish accurate eye-blink detection, we mounted two ArduCAM 2MP Plus cameras directly onto a BiPAP mask using 3D-printed mounts (Fig. 7).

The mount was designed using Autodesk Fusion 360 CAD Software, as this allowed for iterative refinement on the design to ensure that there was optimal camera alignment with patients' eyes. The CAD design focused on stability, accuracy, and patient comfort. Each ArduCAM camera was positioned to capture one eye individually, enabling precise blink detection without requiring facial landmark tracking.. This will help overcome the traditional issues associated with existing eye-tracking devices that rely on capturing the full face, often compromised by the occlusion of the eyes when the patients are wearing respiratory masks.

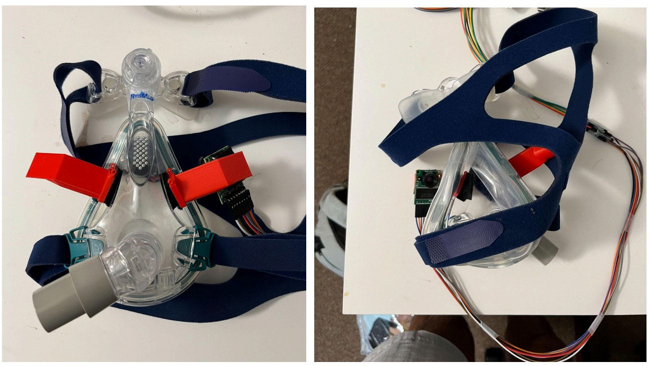

Figure 7. This is the front and back of the physical apparatus with the mounts seen in the front and the wired camera seen on the left side in the image from the back.

Once designed, the mounts were 3D printed with PLA filament. The mounts were secured to the BiPAP mask using adjustable Velcro strips, which provided stability while maintaining comfort and ease of removal for maintenance or adjustments. The cameras were attached to these mounts to ensure minimal vibration or shifting during use, which allowed for a stable video stream.

The ArduCAM cameras are connected via wired connection to a laptop which is running the CNN algorithm. These cameras deliver video feeds at approximately 8 fps, a rate sufficient to detect intentional long blinks (VBPs); however, it may not reliably capture spontaneous blinks. Future improvements could involve optimizing camera settings or hardware to achieve higher frame rates for enhanced algorithm performance.

Ultimately, this camera and mount design provides a stable and unobstructed view of the patient's eyes. This product is a step towards improving ALS patient communication via eye blinks, allowing for greater patient autonomy and a reduction in caregiver dependence.

IRB

An IRB full board review study is currently in the approval process. The review process will resume next academic year (September 2025). The planned protocol is detailed in the Materials and Methods section.

Discussion

General Findings

The result of this project was the development of a high accuracy (90-96%) blink detection algorithm that is able to interface with an ArduCam that is mounted onto a BiPAP mask for an easy-to-use emergency communication device. This device, while not a replacement for more advanced AACs, slots in as an option for patients to use when access to their standard AACs is limited or otherwise untenable. Furthermore, the device's attachment to the patient's BiPAP mask makes for a ubiquitous communication potential.

Limitations and Improvements

Despite its lightweight frame and extensive documentation, the ArduCam Mini 2MP Plus chosen for this project did come with significant drawbacks primarily in the form of its achievable fps. In total and maximum of 8 fps were able to be recorded when the device was running and while this is a serviceable number, it does engender an amount of uncertainty. The camera is able to pick up upon VBPs as they occur in real time, but with only 8 fps there is a risk of missing out on crucial information between frames and increasing the latency of the overall process. Furthermore, as a result of the ArduCam interfacing directly with an Arduino UNO, there was a concern for the connectivity of the wires. Especially when the wires were short, they would often have an incomplete connection with the microcontroller and result in a glitchy/unparsable video stream.

As mentioned previously, switching to 0.5 meter wires incurred a noticeable change in image quality as it allowed the ArduCam a greater range of motion which much less frequently caused the wires to partially disconnect from the microcontroller. However, in total a new hardware setup is likely the answer. An fps of 8 is not worth settling for especially when the ability to reach frames per second on the order of 20-60 is feasible. These two changes together would likely drastically reduce any issues with the device during runtime and improve its overall performance. Additionally, soldering the wires to the board and camera would allow for higher stability to the connection between the wires and camera, as the instability caused corruption with the image transfer.

Furthermore, obtaining a wide variety of image data for training was made impossible through the lack of an IRB for this project. The intended dataset process was to acquire an IRB-HSR and use both subjects without ALS and subjects with ALS to build a dataset of right and left eye images, both closed and open, to conduct training and validation on. This would afford the dataset a wide variety of RGB images from which it could learn over different skin tones, eye colors, and real world data; as opposed to a dataset based on black and white images with a smaller and less diverse dataset learned via transfer learning.

Collecting image data from real subjects, both without and with ALS, on the order of thousands would further help to bolster the predictive capabilities of its device. Furthermore, receiving feedback from subjects with ALS would help to drive the development of the device in a direction which is informed by the people who it is ultimately intended for.

Finally, considering alternative methods for classifying the state of the eye via CNN is worth investigating. This project utilized a CNN trained for binary classification, however papers within the literature use EAR as a matter of thresholding the distance between landmarks around the eyes to determine whether a blink has occurred or not (Dewi et al., 2022). CNNs trained on recognizing landmarks around the eye, as opposed to recognizing whether the eye is closed or open wholesale, have reported similar accuracy results to the ones achieved via this model. As such, it is unclear whether one method of identification would be better suited for this task and experimentation would be necessary.

Materials and Methods

ArduCam

The specific camera used within this device was the ArduCam Mini 2MP Plus OV2640 SPI Camera Module. This camera was chosen primarily for its ease of use and size. Serial Peripheral Interface (SPI) is a communication protocol used within embedded systems to facilitate data communication between a module and its microcontroller, in this case camera and Arduino. This data can then be sent through the Arduino and into a computer and used for image processing. Furthermore, ArduCam has a library of functions which integrate with its hardware streamlining the process of programming the camera and connecting it to the processing unit. Finally, the size of the ArduCam Mini 2MP is 5.08 x 4.34 x 4.01 cm and it weighed roughly 22.96 grams. The camera is necessary to fix onto the BiPAP mask worn by the patient and as such it being small enough to not impede vision and light enough to not affect the fit or function of the mask is imperative to the designs success.

Furthermore, 0.5 meter long M/F wires were used to connect the ArduCam to the Arduino UNO used as its microcontroller. Wires of this length were required to give enough space between the ArduCam and the microcontroller. Thus, letting the microcontroller set a comfortable distance away from the BiPAP mask, instead of needing to be very close to it to accommodate the short wiring distance.

Attachable Mounts

The camera mounts were made with polylactic acid, a biodegradable plastic, at 15% infill and are visualized within Figure 8. The mounts have a 22mm x 24mm x 4mm base with a 34mm straight extension block into a 25 mm curved extension at a 45° angle. They have further been made small for similar reasons to the camera. A lightweight, un-obstructive design is paramount. Furthermore, the mounts are angled inward to provide the camera with an optimal angle to capture footage of the eye region.

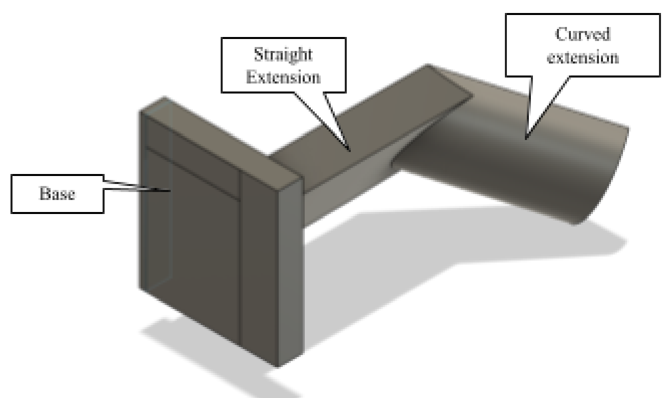

Figure 8. STL file of the 3D printed mount with a platform that was attached to the BiPAP mask with double sided velcro and has an extruded curved piece to allow for camera adjustments to the eye position on the user.

IRB

Subject testing will first occur with healthy individuals, and after some iteration, on ALS patients to minimize burden. Each experiment will start with the subject sitting down on a chair or laying back on a bed and wearing a BiPAP mask that is fitted with the camera apparatus. Each experiment will last 5 to 10 minutes and will be split up into 30 second experimental periods where the subject will either attempt to trigger a VBP or be asked to blink normally. Once the subject has completed the allotted blinking trials, subjects will complete a short questionnaire providing feedback on the device's usability and comfort, administered either verbally or by paper based on subject preference and/or ability.

Author Contributions and Notes

K.D.B and I.S designed the algorithm. D.S. and I.S. designed the camera mount. A.R.N and K.D.B worked on the IRB-HSR study. All authors wrote the paper and poster.

The authors declare no conflict of interest.

Acknowledgments

We would like to thank our project advisors, Dr. Alec Bateman and Dr. Michael DeVore of Barron Associates, Inc., for their consistent guidance and support throughout this project. We also extend our gratitude to Prof. Timothy Allen, the UVA ALS Clinic, and the University of Virginia Department of Biomedical Engineering for their encouragement and contributions to our work.

References

1. Felgoise, S. H., Zaccheo, V., Duff, J., & Simmons, Z. (2016). Verbal communication impacts quality of life in patients with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis & Frontotemporal Degeneration*, *17*(3–4), 179–183. https://doi.org/10.3109/21678421.2015.1125499

2. Brotman, R. G., Moreno-Escobar, M. C., Joseph, J., Munakomi, S., & Pawar, G. (2024). Amyotrophic Lateral Sclerosis. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK556151/

3. Dorst, J., & Ludolph, A. C. (2019). Non-invasive ventilation in amyotrophic lateral sclerosis. *Therapeutic Advances in Neurological Disorders*, *12*, 1756286419857040. https://doi.org/10.1177/1756286419857040

4. Ezzat, M., Maged, M., Gamal, Y., Adel, M., Alrahmawy, M., & El-Metwally, S. (2023). Blink-To-Live eye-based communication system for users with speech impairments. *Scientific Reports*, *13*, 7961. https://doi.org/10.1038/s41598-023-34310-

5. Dewi, C., Chen, R.-C., Jiang, X., & Yu, H. (2022). Adjusting eye aspect ratio for strong eye blink detection based on facial landmarks. *PeerJ Computer Science, 8, e943*. https://doi.org/10.7717/peerj-cs.943

6. Park, S., Spurr, A., & Hilliges, O. (2018). Deep pictorial gaze estimation. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), *Computer Vision – ECCV 2018*. Lecture Notes in Computer Science (Vol. 11211, pp. 741–757). https://doi.org/10.1007/978-3-030-01261-8_44

7. de Wit, J., Bakker, L. A., van Groenestijn, A. C., van den Berg, L. H., Schröder, C. D., Visser-Meily, J. M., & Beelen, A. (2018). Caregiver burden in amyotrophic lateral sclerosis: A systematic review. *Palliative Medicine, 32*(1), 231–245. https://doi.org/10.1177/0269216317709965

8. Jahan, I., Uddin, K. M. A., Murad, S. A., Miah, M. S. U., Khan, T. Z., Masud, M., Aljahdali, S., & Bairagi, A. K. (2023). 4D: A real-time driver drowsiness detector using deep learning. *Electronics, 12*(1), 235. https://doi.org/10.3390/electronics12010235