Article

A Ready-to-Use Grading Tool for Facial Palsy Examiners—Automated Grading System in Facial Palsy Patients Made Easy

Leonard Knoedler [1,*], Maximilian Miragall [2], Martin Kauke-Navarro [3], Doha Obed [4,5], Maximilian Bauer [6], Patrick Tißler [6], Lukas Prantl [1], Hans-Guenther Machens [7], Peter Niclas Broer [8], Helena Baecher [1], Adriana C. Panayi [4], Samuel Knoedler [1,4,7,†] and Andreas Kehrer [1,†]

1   Department of Plastic, Hand and Reconstructive Surgery, University Hospital Regensburg, 93053 Regensburg, Germany
2   Department of Oral and Maxillofacial Surgery, University Hospital Regensburg, 93053 Regensburg, Germany
3   Department of Surgery, Division of Plastic Surgery, Yale New Haven Hospital, Yale School of Medicine, New Haven, CT 06510, USA
4   Department of Surgery, Division of Plastic Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
5   Department of Plastic, Aesthetic, Hand and Reconstructive Surgery, Hannover Medical School, 30625 Hannover, Germany
6   Faculty of Informatics and Data Science, University of Regensburg, 93053 Regensburg, Germany
7   Department of Plastic Surgery and Hand Surgery, Klinikum Rechts der Isar, Technical University of Munich, 81675 Munich, Germany
8   Department of Plastic, Reconstructive, Hand and Burn Surgery, Bogenhausen Academic Teaching Hospital Munich, 81925 Munich, Germany
*   Correspondence: leonard.knoedler@stud.uni-regensburg.de; Tel.: +49-151-448-249-58
†   These authors contributed equally to this work.

Abstract: Background: The grading process in facial palsy (FP) patients is crucial for time- and cost-effective therapy decision-making. The House-Brackmann scale (HBS) represents the most commonly used classification system in FP diagnostics. This study investigated the benefits of linking machine learning (ML) techniques with the HBS. Methods: Image datasets of 51 patients seen at the Department of Plastic, Hand, and Reconstructive Surgery at the University Hospital Regensburg, Germany, between June 2020 and May 2021, were used to build the neural network. A total of nine facial poses per patient were used to automatically determine the HBS. Results: The algorithm had an accuracy of 98%. The algorithm processed the real patient image series (i.e., nine images per patient) in 112 ms. For optimized accuracy, we found 30 training runs to be the most effective training length. Conclusion: We have developed an easy-to-use, time- and cost-efficient algorithm that provides highly accurate automated grading of FP patient images. In combination with our application, the algorithm may facilitate the FP surgeon's clinical workflow.

Keywords: facial palsy; facial paralysis; House-Brackmann scale; artificial intelligence; deep learning; bell's palsy; smile restoration; facial reanimation; application

1. Introduction

Facial palsy (FP) presents with a varying symptom complex attributable to an array of etiologies [1–5]. FP annually affects up to 53 cases per 100,000 population yielding comparable incidence rates across biological sexes [6–9]. Most FP patients are diagnosed with idiopathic FP (Bell's palsy) followed by trauma, viral infections, and tumors [10,11]. Predisposing factors in FP include, for example, hypertension, diabetes mellitus, inflammatory neural demyelination, and migraine [12–15]. The age classes between 45–55 years are particularly prone to develop FP [16]. The sequelae of FP encompass adverse effects on physical, psychological, and social levels. Due to interrupted or erroneous orchestration of

mimic musculature, FP patients encounter flaccidity or synkinetic facial mass movements, respectively [17,18]. Micro- and macroanatomical studies have identified key muscles in FP pathology, such as the depressor anguli oris (DAO), the depressor labii inferioris (DLI), and the zygomaticus major muscles [19–23]. The malfunction of such muscular cornerstones leads to a disfiguring facial appearance and dysfunctional mimic movements [10,24]. Emotional expressiveness is hindered and smile symmetry is impaired [5,25]. The pathognomonic attributes of FP catalyze the manifestation of psychosocial disorders, including anxiety and depression [26]. Tseng et al. demonstrated that FP patients were 59% more likely to develop an anxiety disorder, as compared to unaffected individuals [27]. A 2016 South Korean study found that 32% of FP cases experienced $\geq$2 weeks of depressed mood versus 13% in the general population [28]. Further, increased levels of distress have been observed in FP patients [29]. In a vicious circle, such conditions promote social withdrawal and isolation as well as reduced quality of life [30].

Given the heterogeneous etiology and pathology of FP, only a few general recommendations in FP therapy with a sufficient body of evidence exist. For example, studies recommend the prescription of oral steroids to target acute FP cases [31–33]. The surgical management of FP symptoms ranges from free versus regional muscle transfer to (micro-)surgical techniques, including direct neurorrhaphy and neurotization procedures [34]. For specific indications, even further complex reconstructions have been proposed. Boahene et al. popularized the concept of multivectoral muscle flaps to account for specific human smile pattern, while Klebuc et al. described the DAO-DLI-transfer to address a hypertonic DAO in conjunction with a hypofunctional DLI [35,36]. Azizzadeh et al. have underscored the beneficial effects of modified selective neurectomies to address synkinetic facial musculature counteracting the natural smile [17]. If a patient's eligibility for each surgical technique is critically reviewed and tailored on a case-by-case basis, FP surgery may pave the way for sustainable outcomes.

In each FP case, the grading of the disease severity is crucial to launch appropriate treatment strategies early on and evaluate the course of the FP in follow-up visits. Introduced to the FP community in 1985, the House-Brackmann scale (HBS) has been representing the standard classification system in FP diagnostics across different (non-)surgical specialties [37–40]. The overlaying of evidence-based clinical grading systems and state-of-the-art electronic facial recognition software carries promising potential for objective classification of FP disease [41,42]. However, there is a scarcity of step-by-step tutorials outlining the concrete steps that enable FP surgeons to successfully apply machine learning (ML) techniques in their patient work. We, therefore, aimed to develop an automated facial palsy grading system for FP surgeons interested in ML.

2. Materials and Methods

2.1. Data Acquisition from Facial Patients

From June 2020 to May 2021, prospective data acquisition was performed on 51 patients and additional 10 healthy patients as a control group seen at the Department of Plastic Surgery at the University Hospital Regensburg, Germany (Figure 1).

Inclusion criteria comprised a pathological HBS (i.e., >I) [40]. Of note, the HBS classifies FP severity levels from I (i.e., normal facial function) to VI (i.e., complete FP). Classification is conducted utilizing nine facial expressions (i.e., face in repose; raising the eyebrows; smile with mouth closed; full-denture smile; pursing the lips; gentle eye-closure; forced eye-closure; wrinkling the nose; depressing the lower lip). Facial expressions were recorded based on previous work by Volk and Hadlock [43,44]. As recommended by the Jena facial palsy group, patients were asked to perform these expressions to the best of their ability three times prior to photo documentation [43]. Photo documentation was conducted by either the first or last author (L.K., A.K.) during the last author's facial palsy consultation hours utilizing the CANON EOS 400D with the respective flash unit (Canon, Ota, Japan). The examiner who did not take the patient photos supervised the documentation process. Prior to our first patient photo documentation, we consulted the clinical-intern photo department to evaluate our camera/photography settings. All patient

photos were taken in the same examination room at the same spot to ensure a standardized camera distance. We further used a camera tripod with fixed setting sizes for standardized documentation. In cases in which patients were unable to perform the movement, the authors photographed the best attempt. In cases in which patients stated that they were not used to this facial movement and did not know how to perform the movement, the authors provided the same short instruction on how to theoretically perform the respective movement throughout all cases.
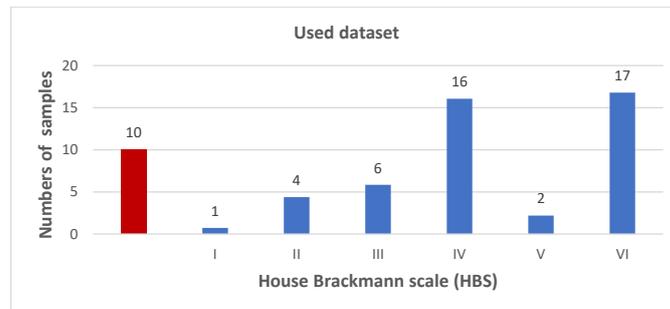

Figure 1. Overview of the study population and distribution of the House-Brackmann scale (HBS). The red bar is visualizing ten healthy individuals as a control group. Facial palsy (FP) patients with HBS scores of IV and VI accounted for the majority of cases, respectively.

We included 51 patients and could therefore validate the network with ten patients since the dataset was divided into a training group with 41 patients and a validation group with ten patients. Of note, there is a difference between the ten patients with FP who were selected from the training data including 51 FP patients by means of a train-test-split and the ten healthy patients who were used for the final validation. The training/validation workflow is illustrated in Figure 2.
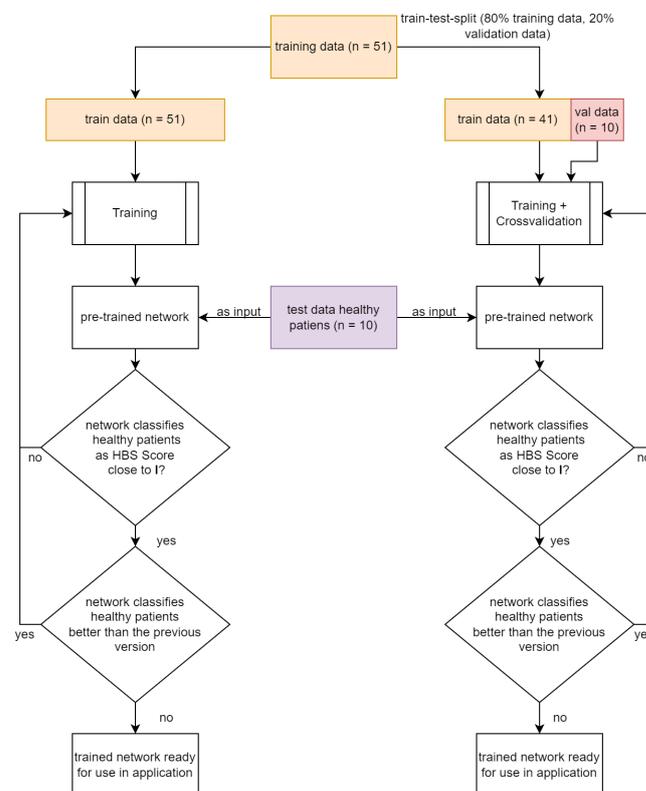

Figure 2. Flow chart of the workflow of how to train the network (training with and without cross-validation) with the training data and validate with healthy patients.

2.2. Facial Palsy Image Segmentation

We designed a facial palsy (FP) image segmentation method as the preprocessing section of the House-Brackmann score classifier, to automatically combine nine input images into one image. Each single image represents a certain facial expression. The nine images serve as input for the neural network, while the House-Brackmann scale (HBS) represented the output value of the network. Beforehand, the images had been pre-classified accordingly by three physicians specialized in FP therapy to set up a distinct link between the nine images and the corresponding HBS. The workflow is illustrated in Figure 2.

Due to the enhanced accuracy of the neural network, and with regard to its possible application in clinical situations, six individual outputs were chosen, each representing one distinct level in the HBS. First, the nine different patient images were implemented in a black–white format and scaled to 200 × 200 pixels to rationalize the computationally intensive training of the neural network. To adapt the nine colored patient images of arbitrary resolution to these requirements, an algorithm was utilized. The mesh yielded 200 × 1800 (Figure 3). The second step is the transformation of nine single pictures to single-composed picture input signals corresponding to the pixels of the nine patient input images and the six output signals, each representing one distinct level in the HBS. Concerning the output signals, each could either have a value of zero or one. For example, an HBS = VI should result in the output value = 1 for the VI. signal, whereas the output values = 0 for the I.–V. signals (Figure 4).
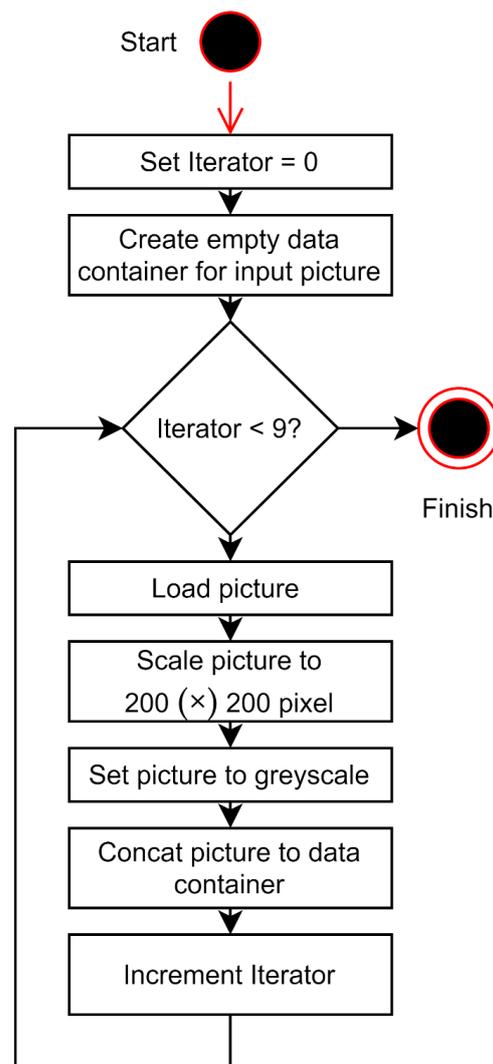

Figure 3. Preliminary image preparation. Transformation of nine single pictures to a single composed picture.
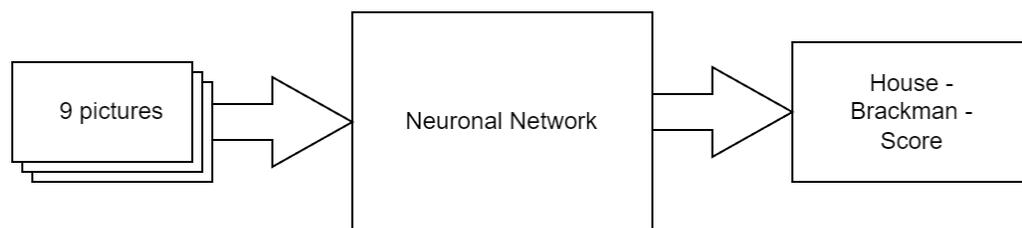
Figure 4. Basic workflow steps. Each patient image was assigned a distinct House-Brackmann scale (HBS) before being added to the neural network.

The neural network training comprises a set of patient images assigned with the corresponding HBS. Each row in the training set, therefore, corresponds to one patient. For training purposes, the data was stored in two arrays with one array for the input and one array for the output data [45].

2.3. Structure of the HBS Score Classifier

For the inner structure of the network, a multi-layer network with three parts was employed ((Figure 5) using machine learning models I, II, and III). The first two layers consist of a convolutional layer, an activation layer including the activation function "relu", and a max-pooling layer. A convolutional layer is a layer in which several neurons are addressed. This enables a more general evaluation of inserted information. This layer can recognize and extract individual features from the input data [46]. A max-pooling layer is used to reduce the computational workload to allow for more efficient processing. Groups of inputs are mapped to individual neurons of the max pooling layer [47]. The activation function "relu" corresponds to the following equation:

$$f(x) = \begin{cases} 0 \; if \; x < 0 \\ x \; if \; x \; \geq 0 \end{cases}$$
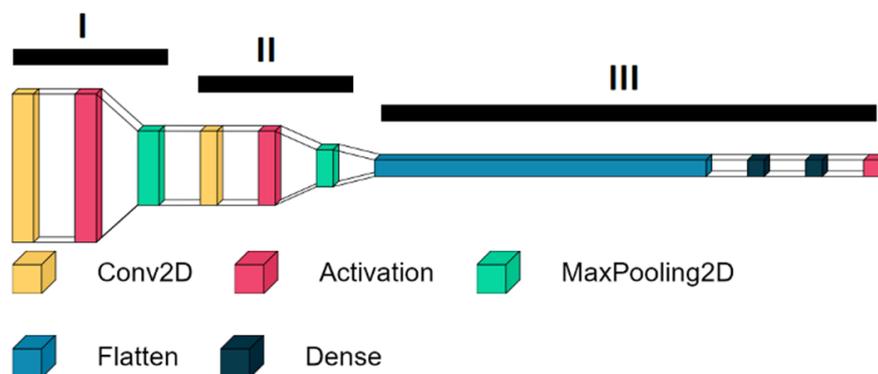

Figure 5. The different components of the machine learning model. The network is subdivided into three parts (i.e., I, II, III).

This function is resource efficient and therefore matches the high throughput of data at the starting point of the neural network.

The classification process is conducted within the convolutional layer and the activation layer, while the max-pooling layer further refines the output, saves computing time, and prevents overfitting by excluding insufficient results. Overfitting leads to an overfitting neural network and occurs when the neural network is trained for too long with the training data, and therefore noise and random outliers in the training data are also adopted as a concept of the model. The problem is that such a trained network can no longer predict new data unknown to it. The size of the three stages is getting continuously smaller in the direction of the output. The output of the second stage is then filtered by a layer of flattening, which connects the second stage with the last stage. The last stage

consists of layers with 64 and six neurons, respectively, with each neuron assigned to a distinct level of the HBS. At the end of the classification process, there is an activation layer including the activation function "sigmoid," which corresponds to the following equation:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Since the results of this function can be between zero or one, this equation is commonly used as a transfer function in the output layer of neural network models to predict probabilities between 0–100%.

For training purposes, 80% of the patient data was used to train the network and the remaining 20% was utilized to validate the neural network. This is called cross-validation. The network underwent varying numbers of training epochs. During each epoch, stochastic gradient descent is used to best configure the neural network to map the input data (i.e., the patient images) and the output data (i.e., the predicted HBS). Following each training run, the network was retested to assess its prediction performance on previously unknown patient data.

Computer operations were performed in the Python programming language (version 3.10.2; Python Software Foundation, Beaverton, OR 97008, USA) on a Lenovo Thinkpad computer (T470, Intel Core i7–7600U processor running at 2.8 GHz with 32 GB of RAM and a Nvidia GeForce GTX 1650 Ti graphic card; Lenovo Deutschland GmbH, 70563 Stuttgart, Germany).

3. Results

Number of Training Runs Determines Prediction Accuracy

Regarding the accuracy rate, 30 training runs proved to be the most effective. The average time of each training run was 9.6 h on our test machine.

The performance of a neural network can be determined using the loss function. This is calculated as follows:

$$L(a_i, y_i) = -(y_i \log(a_i) + (1 - y_i) \log(1 - a_i))$$

In this case, the loss function is used for binary classification, so the output can be zero or one. More precisely, one speaks of the "binary cross entropy loss" function. The index i always refers to the training examples. In the corresponding application, the network was trained with 51 patients and nine images were used to validate the network. The index i is therefore 51. Since it is a binary function, the result can only be zero or one. This calculation then leads to the loss or validation loss of the trained neural network.

After training the network, we had a loss of 0.49 for the training data and a loss of about 0.83 for the validation data. The accuracy for the training data and the validation data was 80% and 52%, respectively.

When training without validation, i.e., using all available patient images without using cross-validation, an accuracy of about 98% was achieved with a loss of less than 0.1. This showed that a longer training of >100 epochs was necessary. After training, the algorithm processed the real patient image series (i.e., nine images per patient) in 112 ms.

Overall performance could be improved by using more training data. Another point of leverage includes adapting the network architecture. To this end, more layers could be added. Further, the resolution of the input data (currently 200×1800 pixels) could be increased. This would render the prediction more independent of physical characteristics, such as beard growth or skin color, which can currently still impair algorithm predictions. Ideally, patients should be asked to remove any coverings, such as hair and/or any other body modification prior to photographic documentation. Another optimization method involves deepening the network structure. Currently, the network consists of three calculation levels, while more calculation levels could be integrated here. The use of non-sequential neural networks (i.e., the insertion of parallel computation strands into the network) can also enhance network performance. This approach is based on the concept that the network

can then simultaneously compute different tasks with different resolutions, meaning that it can detect different templates in the input data.

To test the trained network, data from a healthy control group was used. As the network was only trained with FP patients, the results were expected to be close to an HBS of one. Ten healthy individuals were used as a test group. The results of the control group are shown in Figure 6. Only one individual was assigned a pathological HBS score (i.e., HBS > I) resulting in a false positive rate of 10%.
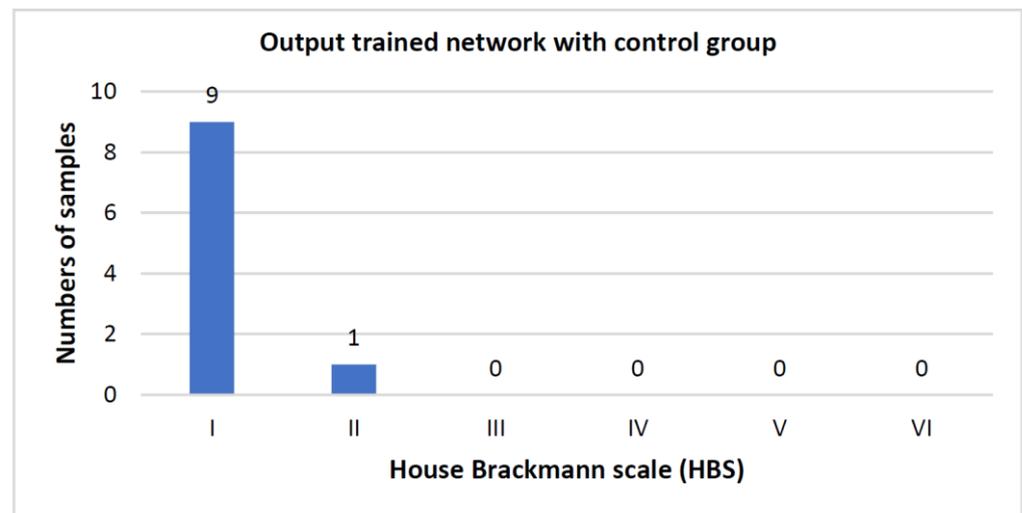
Figure 6. Evaluation of the control group. The control group comprised of ten healthy individuals of whom only one was assigned a pathological House-Brackmann scale (HBS) score (i.e., HBS > I).

To visualize the results of the neural network, an application was coded that implemented different states (i.e., "Init", "Waiting", "Ready", "Error", and "Run"). The workflow of the application is summarized in Video S1.

First, the trained neural network is loaded in the "Init" state. When the nine patient images with the correct coding for the corresponding nine facial expressions are not completely available in the selected folder, the program switches to the "Error" state. The user can return to the "Waiting" state by selecting a correctly filled folder and then proceed to the "Ready" state in which the images are processed according to the aforementioned settings (i.e., black–white format; 200 × 200 pixel resolution). In the "Run" state, the processed images encounter the mesh. The output of the network is displayed as a bar chart. Here, each bar corresponds to the output value of each output neuron of the network (Video S1). Figure 7 illustrates the process workflow of the application.
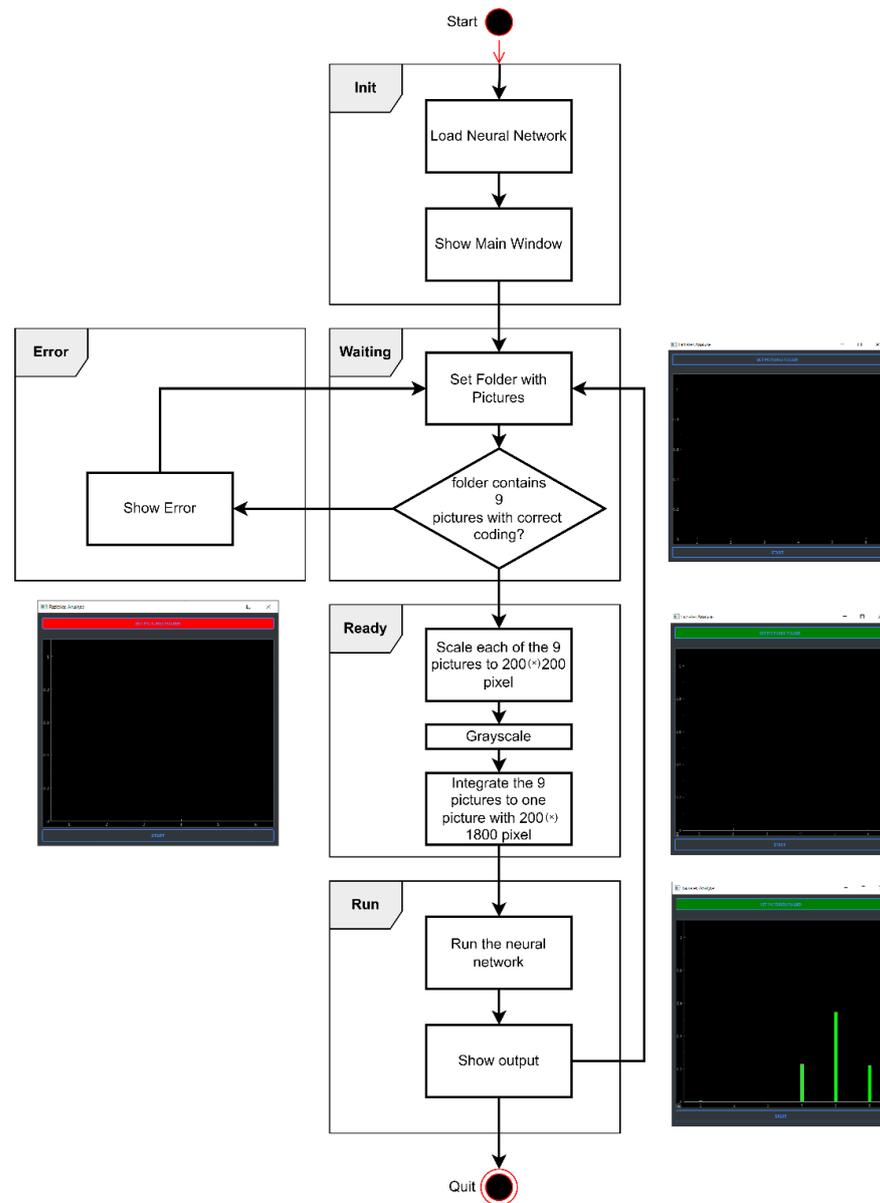
Figure 7. Exemplary application run. Sample output of the network utilizing the application.

4. Discussion

The ever-increasing challenging work environment has resulted in one-third of reconstructive surgeons and surgery residencies reporting burnout symptoms [48]. Yet, recent studies have predicted a future shortage of 3000 US reconstructive surgeons by 2050 and calculated that about 25 million US people have insufficient access to reconstructive surgery services, meaning that a decimated surgery workforce will soon face an increasing work volume [49,50]. This exemplary discrepancy underscores the relevance of time- and cost-efficient tools that facilitate the FP surgeon's workflow. ML has demonstrated beneficial effects in clinical applications, such as in the postoperative monitoring of free flap viability based on skin color or the identification of melanomas using smartphone images [51]. In this study, we provide a time-efficient, user-friendly, and cost-free FP grading algorithm.

In the senior author's experience, thorough grading of FP patients based on the most commonly used classification system, the HBS, can take up to five minutes or even longer in complex FP patient subsets (e.g., neurofibromatosis or apoplex patients). It is not unusual for FP specialists to examine 30–40 FP patients per day, which might accumulate to several hours of grading per day. While these numbers represent worst-case scenarios, the time-

saving potential of automated FP grading is indisputable. Further, additional diagnostic tools, such as ultrasound imaging, have gained popularity in FP examination [52–54]. To include such diagnostic add-ons into the packed clinical routine, FP surgeons first must save time on other tasks such as FP grading. Utilizing our algorithm, we could process real patient image series (i.e., nine images per patient) in 112 ms, on average, which is comparable to the elegant approach developed by Haase et al. (108 ms) [55]. Our model requires only nine standardized patient images, whereas comparable systems have to be fed with video content longer than 20 min per patient [56]. Given the structured simplicity of our model, the entire grading process could be assigned to technical assistants, saving the FP surgeon additional work time and allowing for more time spent on direct patient-doctor communication which has been shown to decrease decisional conflicts and preoperative anxiety from the patient's side [57]. Morrell et al. demonstrated that even five minutes of extra doctor-patient time significantly improved patient satisfaction with their medical provider [58]. From the surgeon's side, such patient-doctor interaction can counteract burnout symptoms and promote work satisfaction [59]. More precisely, repetitive and routine tasks, such as systematic grading, have been identified as burnout drivers, including the recommendation of experts to outsource such work to robotic/computerized assistance tools [60]. Our algorithm may allow for a more refined and self-defined time allocation among the FP surgery workforce.

Recent efforts have focused on combining ML and 3D-frameworks to detect, for example, volume deficits caused by long-term facial musculature atrophy in FP patients [61]. By implementing such techniques, providers aim for advanced grading, ultimately leading to a more differentiated decision-making process in FP therapy [62]. The link between ML and 3D-techniques has resulted in the development of different networks such as AlexNet. Since its launch in 2012, AlexNet has been successfully used in a broad medical application field (e.g., to detect pathologic MRI brain scans or to classify chest X-rays of COVID-19 patients) [63–65]. Based on the HBS, Storey et al. have programmed the 3DPalsyNet, which yielded a classification accuracy of up to 86% (vs. up to 99% in our model). Their algorithm had poor accuracy levels when grading more difficult FP images [66]. Other comparable networks have shown accuracy scores ranging from 88 to 97% [64,66–68]. Zhao et al. demonstrated the prognostic value of a 3D dynamic quantitative analysis system in acute FP cases. However, for each case, the examiner must position six cameras in front of the patient so that every reflective point on the patient's face is detected by at least three cameras [69]. Such preliminary work increases the overall examination time per patient, whereas our platform demonstrated accuracy levels of 99% on images taken with a standard camera widely available in the hospital setting. Anecdotally, the set-up and positioning did not take longer than one minute for our model. Of note, our network can also process images taken with modern smartphones, which may further promote cost-effectiveness. The concept of 3D-technology linked to ML is intriguing, although consequent advantages of such joint systems in grading accuracy when compared to 2D-based platforms remain to be ascertained. Due to their complex and multi-layer neural architecture, such platforms require an extensive and cost-intensive hardware fundament as well as maintenance and acquisition costs of up to $49,000 [70,71]. Advanced programming skills far beyond the FP surgeon's scope are oftentimes needed to develop (and use) such joint systems [72]. Another study by Jiang et al. also involved a highly precise automated grading concept in FP patients. Their work focused on measuring facial skin microcirculation perfusion distribution in FP patients [67]. The Jena group proposed an FP grading index prediction model by using the eFace grading index, which features 16 ordinal fine-grained grading scales for resting face and facial motions [68,73]. The authors addressed objective FP assessment as a linear regression problem instead of an index classification method given the finely graduated ordinal sub-scales of the eFace-scale. Their dataset included image series of 52 multi-ethnical patients of different ages before and after undergoing a hypoglossal-facial anastomosis. Each image series contained nine standardized images of the patient's frontal face. In a second dataset, they included 28 adult healthy subjects as a control study. The

authors reported a mean absolute error (MAE) of 11% in FP patients versus 12% in the control group. The MAE might be further reduced by enlarging the study sample. They also found that deeper networks, such as ResNet-50, did not provide more suitable features for their application, while containing more parameters than a standard VGG-16 model in case fully connected layers were excluded. The authors further outlined the potential adaptation of this approach to be used in other FP scales, such as the Sunnybrook facial grading system [74]. Another study from the Jena group introduced an automated FP grading system based on the Sunnybrook facial grading system [75]. To this end, the authors used 4572 photographs of 233 patients with unilateral peripheral FP. They reported an intraclass coefficient of 0.35 comparing subjective and objective/automated FP grading. The implementation of the Sunnybrook facial grading system carries high translational potential for clinical use, given the recommendation of the Sir Charles Bell Society to use the Sunnybrook facial grading system as a standard grading system for reporting outcomes of facial nerve disorders [76]. Gaber et al. used the Microsoft's Kinect (v2) for real-time FP grading [77]. Their approach was based on the detection of facial landmarks as 3D coordinates both for resting symmetry and voluntary movements, such as raising eyebrows or smiling. Calculation of the regional facial asymmetry was performed through the ratios of distances between corresponding landmarks and a common reference point on the two sides of the face. They also included gamma correction, as well as eye area and mouth slope features. Their system was tested on healthy individuals and showed promising results, yielding a symmetry index of 98% for the ocular region and 96% to 99% for the oral region. A 2017 study by Guo et al. suggested the use of deep convolutional networks for objective FP grading based on the HBS [78]. The authors addressed the problem of confusing neighboring HBS degrees by refining the GoogLeNet model resulting in a classification accuracy of 91% for predicting the HBS degrees. Their dataset included 105 FP subjects versus 75 healthy subjects. Each image set contained four different facial expressions totaling 720 labeled images. Interestingly, the authors designed a data augmentation step to account for the imbalance in HBS degree distribution. Data augmentation included horizontal flipping, random rotating, and resizing, as well as adding salt and pepper noise.

We propose a simple, yet easy-to-use application that allows FP surgeons with varying informatic knowledge to directly utilize our model. With the recent advancements in 3D-technology being promising, we are looking forward to including this innovative technique into our model, as soon as the barriers of cost-effectiveness, user-friendliness, and time-consuming preliminary work have been overcome. Together with other imaging techniques, such as ultrasound or MRI, this approach might enlarge the FP surgeon's diagnostic arsenal and allow for comprehensive patient evaluation at different time points of FP therapy (Figure 8) [54].
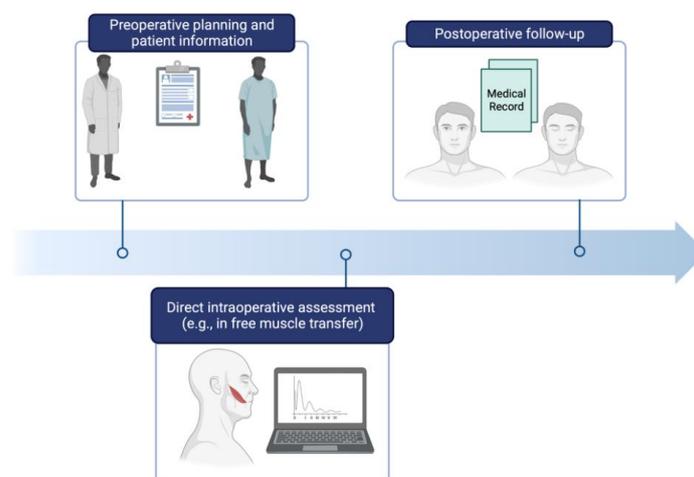

Figure 8. Implementation of automated grading in the clinical workflow. Automated grading could be used in the preoperative planning phase, as well as for direct intraoperative assessment. Following (non-)surgical therapy, automated grading may allow for standardizing patient follow-up evaluation.

Limitations

The present study is not without limitations. Our study population comprised a disproportionate percentage of severe FP cases. To account for this imbalance, we performed oversampling. We included 51 patients in this study. Therefore, large-scale studies are needed to corroborate our findings and demonstrate the efficiency of our algorithm in larger patient cohorts. However, our study population did accurately represent the most common clinical FP scenarios. The HBS represents the most commonly used FP grading classification system in US clinics but has revealed certain downsides such as the insufficient implementation of synkinesis [79]. Thus, we aim to translate the algorithm into more sophisticated grading systems, such as those developed by Guarin and Hadlock [41,80,81]. Work done by the Jena group underscored the implementability of automated grading approaches into the Sunnybrook facial grading system [75]. The study by Guo et al. provided further potential points of leverage to target the imbalance of HBS degree distribution [78], while our study demonstrated the general feasibility of combining all photos to generate one single score. Yet, further efforts are needed toward creating a tensor with the nine images per FP patient instead of combining the images which can cause dilution of the information present in the images.

5. Conclusions

We have developed an easy-to-use, time- and cost-efficient, as well as highly accurate algorithm utilizing ML principles. Integrated into a user-friendly application, our model may facilitate and accelerate the FP surgeon's clinical workflow.

References

1. Jowett, N. A General Approach to Facial Palsy. *Otolaryngol. Clin. N. Am.* **2018**, *51*, 1019–1031. [CrossRef]
2. Teresa, M.O.; Jowett, N.; Hadlock, T.A. Facial Palsy: Diagnostic and Therapeutic Management. *Otolaryngol. Clin. N. Am.* **2018**, *51*, xvii–xviii.
3. Owusu, J.A.; Stewart, C.M.; Boahene, K. Facial Nerve Paralysis. *Med. Clin. N. Am.* **2018**, *102*, 1135–1143. [CrossRef] [PubMed]
4. McCaul, J.A.; Cascarini, L.; Godden, D.; Coombes, D.; Brennan, P.A.; Kerawala, C.J. Evidence based management of Bell's palsy. *Br. J. Oral Maxillofac. Surg.* **2014**, *52*, 387–391. [CrossRef] [PubMed]
5. Kosins, A.M.; Hurvitz, K.A.; Evans, G.R.; Wirth, G.A. Facial paralysis for the plastic surgeon. *Can. J. Plast. Surg.* **2007**, *15*, 77–82. [CrossRef] [PubMed]
6. Yanagihara, N. Incidence of Bell's palsy. *Ann. Otol. Rhinol. Laryngol. Suppl.* **1988**, *137*, 3–4. [CrossRef]
7. Bleicher, J.N.; Hamiel, S.; Gengler, J.S.; Antimarino, J. A survey of facial paralysis: Etiology and incidence. *Ear Nose Throat J.* **1996**, *75*, 355–358. [CrossRef]
8. Heckmann, J.G.; Urban, P.P.; Pitz, S.; Guntinas-Lichius, O.; Gágyor, I. The Diagnosis and Treatment of Idiopathic Facial Paresis (Bell's Palsy). *Dtsch. Arztebl. Int.* **2019**, *116*, 692–702. [CrossRef]

9. Zhang, W.; Xu, L.; Luo, T.; Wu, F.; Zhao, B.; Li, X. The etiology of Bell's palsy: A review. *J. Neurol.* **2020**, *267*, 1896–1905. [CrossRef]
10. Roob, G.; Fazekas, F.; Hartung, H.P. Peripheral facial palsy: Etiology, diagnosis and treatment. *Eur. Neurol.* **1999**, *41*, 3–9. [CrossRef]
11. Tiemstra, J.D.; Khatkhate, N. Bell's palsy: Diagnosis and management. *Am. Fam. Physician* **2007**, *76*, 997–1002. [PubMed]
12. Abraham-Inpijn, L.; Devriese, P.P.; Hart, A.A. Predisposing factors in Bell's palsy: A clinical study with reference to diabetes mellitus, hypertension, clotting mechanism and lipid disturbance. *Clin. Otolaryngol. Allied. Sci.* **1982**, *7*, 99–105. [CrossRef] [PubMed]
13. Greco, D.; Gambina, F.; Pisciotta, M.; Abrignani, M.; Maggio, F. Clinical characteristics and associated comorbidities in diabetic patients with cranial nerve palsies. *J. Endocrinol. Investig.* **2012**, *35*, 146–149.
14. Liston, S.L.; Kleid, M.S. Histopathology of Bell's palsy. *Laryngoscope* **1989**, *99*, 23–26. [CrossRef] [PubMed]
15. Peng, K.P.; Chen, Y.T.; Fuh, J.L.; Tang, C.H.; Wang, S.J. Increased risk of Bell palsy in patients with migraine: A nationwide cohort study. *Neurology* **2015**, *84*, 116–124. [CrossRef]
16. Hohman, M.H.; Hadlock, T.A. Etiology, diagnosis, and management of facial palsy: 2000 patients at a facial nerve center. *Laryngoscope* **2014**, *124*, E283–E293. [CrossRef]
17. Azizzadeh, B.; Irvine, L.E.; Diels, J.; Slattery, W.H.; Massry, G.G.; Larian, B.; Riedler, K.L.; Peng, G.L. Modified Selective Neurectomy for the Treatment of Post-Facial Paralysis Synkinesis. *Plast. Reconstr. Surg.* **2019**, *143*, 1483–1496. [CrossRef]
18. Biglioli, F.; Kutanovaite, O.; Rabbiosi, D.; Colletti, G.; Mohammed, M.; Saibene, A.M.; Cupello, S.; Privitera, A.; Battista, V.M.; Lozza, A.; et al. Surgical treatment of synkinesis between smiling and eyelid closure. *J. Craniomaxillofac Surg.* **2017**, *45*, 1996–2001. [CrossRef]
19. Kehrer, A.; Engelmann, S.; Ruewe, M.; Geis, C.; Taeger, C.; Kehrer, M.; Prantl, L.; Tamm, E.; Bleys, R.R.L.A.W.; Mandlik, V. Anatomical study of the zygomatic and buccal branches of the facial nerve: Application to facial reanimation procedures. *Clin. Anat.* **2019**, *32*, 480–488. [CrossRef]
20. Kehrer, A.; Engelmann, S.; Bauer, R.; Taeger, C.; Grechenig, S.; Kehrer, M.; Prantl, L.; Tamm, E.R.; Bleys, R.R.L.A.W.; Mandlik, V. The nerve supply of zygomaticus major: Variability and distinguishing zygomatic from buccal facial nerve branches. *Clin. Anat.* **2018**, *31*, 560–565. [CrossRef] [PubMed]
21. D'Andrea, E.; Barbaix, E. Anatomic research on the perioral muscles, functional matrix of the maxillary and mandibular bones. *Surg. Radiol. Anat.* **2006**, *28*, 261–266. [CrossRef] [PubMed]
22. Engelmann, S.; Ruewe, M.; Geis, S.; Taeger, C.D.; Kehrer, M.; Tamm, E.R.; Bleys, R.L.A.W.; Zeman, F.; Prantl, L.; Kehrer, A. Rapid and Precise Semi-Automatic Axon Quantification in Human Peripheral Nerves. *Sci. Rep.* **2020**, *10*, 1935. [CrossRef] [PubMed]
23. Mandlik, V.; Ruewe, M.; Engelmann, S.; Geis, S.; Taeger, C.; Kehrer, M.; Tamm, E.R.; Bleys, R.; Prantl, L.; Kehrer, A. Significance of the Marginal Mandibular Branch in Relation to Facial Palsy Reconstruction: Assessment of Microanatomy and Macroanatomy Including Axonal Load in 96 Facial Halves. *Ann. Plast. Surg.* **2019**, *83*, e43–e49. [CrossRef] [PubMed]
24. Toulgoat, F.; Sarrazin, J.; Benoudiba, F.; Pereon, Y.; Auffray-Calvier, E.; Daumas-Duport, B.; Lintia-Gaultier, A.; Desal, H. Facial nerve: From anatomy to pathology. *Diagn. Interv. Imaging* **2013**, *94*, 1033–1042. [CrossRef] [PubMed]
25. Natghian, H.; Fransen, J.; Rozen, S.M.; Rodriguez-Lorenzo, A. Qualitative and Quantitative Analysis of Smile Excursion in Facial Reanimation: A Systematic Review and Meta-analysis of 1- versus 2-stage Procedures. *Plast. Reconstr. Surg. Glob. Open* **2017**, *5*, e1621. [CrossRef]
26. Dobel, C.; Miltner, W.H.; Witte, O.W.; Volk, G.F.; Guntinas-Lichius, O. Emotional impact of facial palsy. *Laryngorhinootologie* **2013**, *92*, 9–23.
27. Tseng, C.C.; Hu, L.Y.; Liu, M.E.; Yang, A.C.; Shen, C.C.; Tsai, S.J. Bidirectional association between Bell's palsy and anxiety disorders: A nationwide population-based retrospective cohort study. *J. Affect. Disord.* **2017**, *215*, 269–273. [CrossRef]
28. Chang, Y.-S.; Choi, J.E.; Kim, S.W.; Baek, S.-Y.; Cho, Y.-S. Prevalence and associated factors of facial palsy and lifestyle characteristics: Data from the Korean National Health and Nutrition Examination Survey 2010–2012. *BMJ Open* **2016**, *6*, e012628. [CrossRef]
29. Hotton, M.; Huggons, E.; Hamlet, C.; Shore, D.; Johnson, D.; Norris, J.H.; Kilcoyne, S.; Dalton, L. The psychosocial impact of facial palsy: A systematic review. *Br. J. Health Psychol.* **2020**, *25*, 695–727. [CrossRef]
30. Coulson, S.E.; O'Dwyer, N.J.; Adams, R.D.; Croxson, G.R. Expression of emotion and quality of life after facial nerve paralysis. *Otol. Neurotol.* **2004**, *25*, 1014–1019. [CrossRef]
31. Ramsey, M.J.; DerSimonian, R.; Holtel, M.R.; Burgess, L.P. Corticosteroid treatment for idiopathic facial nerve paralysis: A meta-analysis. *Laryngoscope* **2000**, *110*, 335–341. [CrossRef] [PubMed]
32. Madhok, V.B.; Gagyor, I.; Daly, F.; Somasundara, D.; Sullivan, M.; Gammie, F.; Sullivan, F. Corticosteroids for Bell's palsy (idiopathic facial paralysis). *Cochrane Database Syst. Rev.* **2016**, *7*, Cd001942. [CrossRef] [PubMed]
33. Quant, E.C.; Jeste, S.S.; Muni, R.H.; Cape, A.V.; Bhussar, M.K.; Peleg, A.Y. The benefits of steroids versus steroids plus antivirals for treatment of Bell's palsy: A meta-analysis. *BMJ* **2009**, *339*, b3354. [CrossRef]
34. Klebuc, M. The evolving role of the masseter-to-facial (V-VII) nerve transfer for rehabilitation of the paralyzed face. *Ann. Chir. Plast. Esthet.* **2015**, *60*, 436–441. [CrossRef]
35. Klebuc, M.J.A.; (Labio-mental Synkinetic Dysfunction Weill Cornell School of Medicine, New York, NY, USA). Personal communication, 2020.

36. Boahene, K.O.; Owusu, J.; Ishii, L.; Ishii, M.; Desai, S.; Kim, I.; Kim, L.; Byrne, P. The Multivector Gracilis Free Functional Muscle Flap for Facial Reanimation. *JAMA Facial Plast. Surg.* **2018**, *20*, 300–306. [CrossRef] [PubMed]

37. Reitzen, S.D.; Babb, J.S.; Lalwani, A.K. Significance and reliability of the House-Brackmann grading system for regional facial nerve function. *Otolaryngol. Head Neck Surg.* **2009**, *140*, 154–158. [CrossRef]

38. Fattah, A.Y.; Gurusinghe, A.D.R.; Gavilan, J.; Hadlock, T.A.; Marcus, J.R.; Marres, H.; Nduka, C.C.; Slattery, W.H.; Snyder-Warwick, A.K.; Sir Charles Bell Society. Facial nerve grading instruments: Systematic review of the literature and suggestion for uniformity. *Plast. Reconstr. Surg.* **2015**, *135*, 569–579. [CrossRef] [PubMed]

39. Sun, M.Z.; Oh, M.C.; Safaee, M.; Kaur, G.; Parsa, A.T. Neuroanatomical correlation of the House-Brackmann grading system in the microsurgical treatment of vestibular schwannoma. *Neurosurg. Focus* **2012**, *33*, E7. [CrossRef] [PubMed]

40. House, J.W.; Brackmann, D.E. Facial nerve grading system. *Otolaryngol. Head Neck Surg.* **1985**, *93*, 146–147. [CrossRef]

41. Dusseldorp, J.R.; van Veen, M.M.; Mohan, S.; Hadlock, T.A. Outcome Tracking in Facial Palsy. *Otolaryngol. Clin. N. Am.* **2018**, *51*, 1033–1050. [CrossRef]

42. Banks, C.A.; Jowett, N.; Azizzadeh, B.; Beurskens, C.; Bhama, P.; Borschel, G.; Coombs, C.; Coulson, S.; Croxon, G.; Diels, J.; et al. Worldwide Testing of the eFACE Facial Nerve Clinician-Graded Scale. *Plast. Reconstr. Surg.* **2017**, *139*, 491e–498e. [CrossRef]

43. Schaede, R.A.; Volk, G.F.; Modersohn, L.; Barth, J.M.; Denzler, J.; Guntinas-Lichius, O. Patienten-Instruktionsvideo mit synchroner Videoaufnahme von Gesichtsbewegungen bei Fazialisparese [Video Instruction for Synchronous Video Recording of Mimic Movement of Patients with Facial Palsy]. *Laryngorhinootologie* **2017**, *96*, 844–849.

44. Santosa, K.B.; Fattah, A.; Gavilán, J.; Hadlock, T.A. Snyder-Warwick AK. Photographic Standards for Patients With Facial Palsy and Recommendations by Members of the Sir Charles Bell Society. *JAMA Facial Plast. Surg.* **2017**, *19*, 275–281. [CrossRef] [PubMed]

45. Perkel, J.M. Programming: Pick up Python. *Nature* **2015**, *518*, 125–126. [CrossRef] [PubMed]

46. Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. *Insights Imaging* **2018**, *9*, 611–629. [CrossRef] [PubMed]

47. Suárez-Paniagua, V.; Segura-Bedmar, I. Evaluation of pooling operations in convolutional architectures for drug-drug interaction extraction. *BMC Bioinform.* **2018**, *19* (Suppl. S8), 209. [CrossRef] [PubMed]

48. Ribeiro, R.V.E.; Martuscelli, O.J.D.; Vieira, A.C.; Vieira, C.F. Prevalence of Burnout among Plastic Surgeons and Residents in Plastic Surgery: A Systematic Literature Review and Meta-analysis. *Plast. Reconstr. Surg. Glob. Open* **2018**, *6*, e1854. [CrossRef] [PubMed]

49. Yang, J.; Jayanti, M.K.; Taylor, A.; Williams, T.E.; Tiwari, P. The impending shortage and cost of training the future plastic surgical workforce. *Ann. Plast. Surg.* **2014**, *72*, 200–203. [CrossRef]

50. Bauder, A.R.; Sarik, J.R.; Butler, P.D.; Noone, R.B.; Fischer, J.P.; Serletti, J.M.; Kanchwala, S.K.; Kovach, S.J.; Fox, J.P. Geographic Variation in Access to Plastic Surgeons. *Ann. Plast. Surg.* **2016**, *76*, 238–243. [CrossRef]

51. Jarvis, T.; Thornburg, D.; Rebecca, A.M.; Teven, C.M. Artificial Intelligence in Plastic Surgery: Current Applications, Future Directions, and Ethical Implications. *Plast. Reconstr. Surg. Glob. Open* **2020**, *8*, e3200. [CrossRef]

52. Sauer, M.; Guntinas-Lichius, O.; Volk, G.F. Ultrasound echomyography of facial muscles in diagnosis and follow-up of facial palsy in children. *Eur. J. Paediatr. Neurol.* **2016**, *20*, 666–670. [CrossRef] [PubMed]

53. Volk, G.F.; Wystub, N.; Pohlmann, M.; Finkensieper, M.; Chalmers, H.J.; Guntinas-Lichius, O. Quantitative ultrasonography of facial muscles. *Muscle Nerve* **2013**, *47*, 878–883. [CrossRef] [PubMed]

54. Kehrer, A.; Ruewe, M.; Klebuc, M.; Platz Batista da Silva, N.; Lonic, D.; Heidkrueger, P.; Jung, E.M.; Prantl, L.; Knoedler, L. Objectifying the Antagonistic Role of the Depressor Anguli Oris Muscle in Synkinetic Smile Formation Utilizing High-Resolution Ultrasound—A Prospective Study. *Plast. Reconstr. Surg.* **2022**, *ahead of print*.

55. Haase, D.; Minnigerode, L.; Volk, G.F.; Denzler, J.; Guntinas-Lichius, O. Automated and objective action coding of facial expressions in patients with acute facial palsy. *Eur. Arch. Oto-Rhino-Laryngol.* **2015**, *272*, 1259–1267. [CrossRef]

56. Meier-Gallati, V.; Scriba, H.; Fisch, U.P. Objective scaling of facial nerve function based on area analysis (OSCAR). *Otolaryngol. Head Neck Surg.* **1998**, *118*, 545–550.

57. Shinkunas, L.A.; Klipowicz, C.J.; Carlisle, E.M. Shared decision making in surgery: A scoping review of patient and surgeon preferences. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 190. [CrossRef] [PubMed]

58. Morrell, D.; Evans, M.; Morris, R.; Roland, M. The "five minute" consultation: Effect of time constraint on clinical content and patient satisfaction. *Br. Med. J. (Clin. Res. Ed)* **1986**, *292*, 870–873. [CrossRef] [PubMed]

59. IsHak, W.W.; Lederer, S.; Mandili, C.; Nikravesh, R.; Seligman, L.; Vasa, M.; Ogunyemi, D.; Bernstein, C.A. Burnout during residency training: A literature review. *J. Grad. Med. Educ.* **2009**, *1*, 236–242. [CrossRef] [PubMed]

60. Edú-Valsania, S.; Laguía, A.; Moriano, J.A. Burnout: A Review of Theory and Measurement. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1780. [CrossRef]

61. Olivetto, M.; Sarhan, F.-R.; Ben Mansour, K.; Marie, J.-P.; Marin, F.; Dakpé, S. Quantitative Analysis of Facial Palsy Based on 3D Motion Capture (SiMoVi—FaceMoCap Project). *Arch. Phys. Med. Rehabil.* **2019**, *100*, e112. [CrossRef]

62. Su, Z.; Liang, B.; Shi, F.; Gelfond, J.; Šegalo, S.; Wang, J.; Jia, P.; Hao, X. Deep learning-based facial image analysis in medical research: A systematic review protocol. *BMJ Open* **2021**, *11*, e047549. [CrossRef] [PubMed]

63. Singh, S.P.; Wang, L.; Gupta, S.; Goli, H.; Padmanabhan, P.; Gulyás, B. 3D Deep Learning on Medical Images: A Review. *Sensors* **2020**, *20*, 5097. [CrossRef] [PubMed]

64. Lu, S.; Wang, S.-H.; Zhang, Y.-D. Detection of abnormal brain in MRI via improved AlexNet and ELM optimized by chaotic bat algorithm. *Neural Comput. Appl.* **2021**, *33*, 10799–10811. [CrossRef]
65. Pérez, E.; Sanchez, S. Deep Learning Transfer with AlexNet for chest X-ray COVID-19 recognition. *IEEE Access* **2020**, *100*, 4336.
66. Storey, G.; Jiang, R.; Keogh, S.; Bouridane, A.; Li, C.-T. 3DPalsyNet: A Facial Palsy Grading and Motion Recognition Framework Using Fully 3D Convolutional Neural Networks. *IEEE Access* **2019**, *7*, 121655–121664. [CrossRef]
67. Jiang, C.; Wu, J.; Zhong, W.; Wei, M.; Tong, J.; Yu, H.; Wang, L. Automatic Facial Paralysis Assessment via Computational Image Analysis. *J. Healthc. Eng.* **2020**, *2020*, 2398542. [CrossRef] [PubMed]
68. Raj, A.; Mothes, O.; Sickert, S.; Volk, G.F.; Guntinas-Lichius, O.; Denzler, J. Automatic and Objective Facial Palsy Grading Index Prediction Using Deep Feature Regression. In *Medical Image Understanding and Analysis, Proceedings of the MIUA 2020 Communications in Computer and Information, Oxford, UK, 15–17 July 2020*; SciencePapież, B., Namburete, A., Yaqub, M., Noble, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; Volume 1248.
69. Zhao, Y.; Feng, G.; Wu, H.; Aodeng, S.; Tian, X.; Volk, G.F.; Guntinas-Lichius, O.; Gao, Z. Prognostic value of a three-dimensional dynamic quantitative analysis system to measure facial motion in acute facial paralysis patients. *Head Face Med.* **2020**, *16*, 15. [CrossRef] [PubMed]
70. Overschmidt, B.; Qureshi, A.A.; Parikh, R.P.; Yan, Y.; Tenenbaum, M.M.; Myckatyn, T.M. A prospective evaluation of three-dimensional image simulation: Patient-reported outcomes and mammometrics in primary breast augmentation. *Plast. Reconstr. Surg.* **2018**, *142*, 133e–144e. [CrossRef]
71. Wesselius, T.S.; Verhulst, A.C.; Vreeken, R.D.; Xi, T.; Maal, T.J.; Ulrich, D.J. Accuracy of three software applications for breast volume calculations from three-dimensional surface images. *Plast. Reconstr. Surg.* **2018**, *142*, 858–865. [CrossRef] [PubMed]
72. Rodríguez-Pérez, R.; Bajorath, J. Interpretation of Compound Activity Predictions from Complex Machine Learning Models Using Local Approximations and Shapley Values. *J. Med. Chem.* **2020**, *63*, 8761–8777. [CrossRef]
73. Guntinas-Lichius, O.; Denzler, J. Automatic and objective facial palsy grading index prediction using deep feature regression. In *Medical Image Understanding and Analysis, Proceedings of the 24th Annual Conference, Geneva, Switzerland, 15 November 2022*; Springer: Berlin/Heidelberg, Germany, 2022.
74. Ross, B.R.; Fradet, G.; Nedzelski, J.M. Development of a sensitive clinical facial grading system. *Eur. Arch. Otorhinolaryngol.* **1996**, *114*, 380–386.
75. Mothes, O.; Modersohn, L.; Volk, G.F.; Klingner, C.; Witte, O.W.; Schlattmann, P.; Denzler, J.; Guntinas-Lichius, O. Automated objective and marker-free facial grading using photographs of patients with facial palsy. *Eur. Arch. Otorhinolaryngol.* **2019**, *276*, 3335–3343. [CrossRef] [PubMed]
76. Fattah, A.Y.; Gavilan, J.; Hadlock, T.A.; Marcus, J.R.; Marres, H.; Nduka, C.; Slattery, W.H.; Snyder-Warwick, A.K. Survey of methods of facial palsy documentation in use by members of the Sir Charles Bell Society. *Laryngoscope* **2014**, *124*, 2247–2251. [CrossRef] [PubMed]
77. Gaber, A.; Taher, M.F.; Wahed, M.A. Quantifying facial paralysis using the Kinect v2. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 2497–2501.
78. Guo, Z.; Shen, M.; Duan, L.; Zhou, Y.; Xiang, J.; Ding, H.; Chen, S.; Deussen, O.; Dan, G. Deep assessment process: Objective assessment process for unilateral peripheral facial paralysis via deep convolutional neural network. In Proceedings of the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), Melbourne, Australia, 18–21 April 2017; pp. 135–138. [CrossRef]
79. Miller, M.Q.; Hadlock, T.A.; Fortier, E.; Guarin, D.L. The Auto-eFACE: Machine Learning-Enhanced Program Yields Automated Facial Palsy Assessment Tool. *Plast. Reconstr. Surg.* **2021**, *147*, 467–474. [CrossRef]
80. Guarin, D.L.; Yunusova, Y.; Taati, B.; Dusseldorp, J.R.; Mohan, S.; Tavares, J.; van Veen, M.M.; Fortier, E.; Hadlock, T.A.; Jowett, N. Toward an Automatic System for Computer-Aided Assessment in Facial Palsy. *Facial Plast. Surg. Aesthet. Med.* **2020**, *22*, 42–49. [CrossRef] [PubMed]
81. Banks, C.A.; Bhama, P.K.; Park, J.; Hadlock, C.R.; Hadlock, T.A. Clinician-Graded Electronic Facial Paralysis Assessment: The eFACE. *Plast. Reconstr. Surg.* **2015**, *136*, 223e–230e. [CrossRef] [PubMed]