Article

A Study on the Possible Diagnosis of Parkinson's Disease on the Basis of Facial Image Analysis

Jacek Jakubowski ^{1,*} , Anna Potulska-Chromik ², Kamila Białek ¹, Monika Nojszewska ²
and Anna Kostera-Pruszczyk ²

¹ Faculty of Electronics, Military University of Technology, 00-908 Warsaw, Poland; kamila.bialek@wat.edu.pl

² Department of Neurology, Medical University of Warsaw, 02-097 Warsaw, Poland;

apotulska@wum.edu.pl (A.P.-C.); monika.nojszewska@wum.edu.pl (M.N.);

anna.kostera-pruszczyk@wum.edu.pl (A.K.-P.)

* Correspondence: jacek.jakubowski@wat.edu.pl; Tel.: +48-261-837-446

Abstract: One of the symptoms of Parkinson's disease is the occurrence of problems with the expression of emotions on the face, called facial masking, facial bradykinesia or hypomimia. Recent medical studies show that this symptom can be used in the diagnosis of this disease. In the presented study, the authors, on the basis of their own research, try to answer the question of whether it is possible to build an automatic Parkinson's disease recognition system based on the face image. The research used image recordings in the field of visible light and infrared. The material for the study consisted of registrations in a group of patients with Parkinson's disease and a group of healthy patients. The patients were asked to express a neutral facial expression and a smile. In the detection, both geometric and holistic methods based on the use of convolutional network and image fusion were used. The obtained results were assessed quantitatively using statistical measures, including F_1 score, which was a value of 0.941. The results were compared with a competitive work on the same subject. A novelty of our experiments is that patients with Parkinson's disease were in the so-called ON phase, in which, due to the action of drugs, the symptoms of the disease are reduced. The results obtained seem to be useful in the process of early diagnosis of this disease, especially in times of remote medical examination.

Keywords: Parkinson's disease; image analysis; convolutional networks

Citation: Jakubowski, J.; Potulska-Chromik, A.; Białek, K.; Nojszewska, M.; Kostera-Pruszczyk, A. A Study on the Possible Diagnosis of Parkinson's Disease on the Basis of Facial Image Analysis. *Electronics* **2021**, *10*, 2832. <https://doi.org/10.3390/electronics10222832>

Academic Editors: Christian Morbidoni, Francesco Di Nardo and Alessandro Cucchiarelli

Received: 22 October 2021

Accepted: 16 November 2021

Published: 18 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Contemporary diagnosis of Parkinson's disease (PD) is based on a direct clinical picture and medical history. Its basis is the finding of the coexistence of the following symptoms: muscular stiffness, tremor of the limbs, slowness and postural disorders. It is estimated that more than 10 million people worldwide suffer from the disease [1], especially after the age of 50 [2]. Making a diagnosis in the case of advanced Parkinson's disease is not difficult for medical doctors and is based on visible clinical signs. The real problem is the correct diagnosis made at an early stage of the disease when the severity of the classic symptoms is small. It is estimated that incorrect diagnosis can occur in about 10–25% of cases [2,3].

The dominant modern medical standard for the diagnosis and assessment of the severity of Parkinson's disease symptoms is the UPDRS (Unified Parkinson's Disease Rating Scale) adopted for use in 1987 [4,5]. This is a patient evaluation scale using four different components in a five-step gradation. Higher scale ranges correspond to more advanced stages of the disease. On the basis of the interview and the tests performed, this scale determines the intellectual state and mood disorders, the quality of everyday life and motor functions. However, the use of the scale requires a good knowledge of the symptoms of the disease and extensive clinical experience of the examining physician. For this reason, nowadays, many scientific centers are making efforts to objectify the tests

carried out as well as efforts to find new biomarkers that could be used to support the process of diagnosis and progression of the disease at an early stage of its development. For a brief overview of the methods being developed, see Section 2.1. The scope of this study was to investigate the possibility of building an automated machine learning system fed with facial images acquired by a camera in order to recognize patients with Parkinson's disease. The key motivation to undertake the research on the use of facial images is the observed difficulty in expressing emotions in PD patients' faces due to the deterioration of the speed and coordination of the action of the relevant muscles—the so-called hypomimia, facial masking or facial bradykinesia [6]. Recent studies showed that it could be considered a very sensitive biomarker for PD as it is highly correlated with other symptoms [7]. In clinical practice, the assessment of facial masking is made personally by a natural visual inspection by a medical doctor during the medical interview. Instead, we proposed using algorithms of machine learning. We investigated the possibilities of making decisions on the basis of face images acquired in the field of visible and infrared imagery. The reason for the additional use of infrared images is two-fold. First, thermal imagery is less dependent on ambient lighting conditions as compared to visible light imagery. Second, the temperature distribution on the surface of the skin may be another useful biomarker in Parkinson's disease. Abnormalities of the skin caused by the deposition of pathological alpha-synuclein in it are seen in many patients [8]. Recent research demonstrated that PD patients had altered skin blood flow regulation and showed abnormal thermal responses of the skin [9].

The main novelty presented in this paper, in contrast to other works devoted to the detection of hypomimia as a PD marker, is the use of data recorded in conditions of low severity of disease symptoms. These conditions were obtained through the use of drugs by patients, namely oral L-DOPA. It was assumed that the reduced severity of the symptoms of the disease corresponds to the early stage of the development of the disease, i.e., when its correct diagnosis is the most difficult.

In the last decade, machine learning algorithms have shown excellent performance in the field of biomedical engineering. They can successfully replace conventional methods of visual inspection, which are very often tedious and error-prone. Recent achievements in the use of machine learning methods, including deep learning for the diagnosis of various diseases in medical images, e.g., brain tumors or skin cancers [10,11], have become a motivation to undertake the research presented in this work.

2. Related Works

2.1. Evaluation of Tremor, Speech, Hand Writing and Gait

Most of the methods developed to support PD diagnostics consist of the assessment of the patient's motor functions based on measurements and analyses of one-dimensional biosignals; their potential in diagnosis results from the etiology of this disease. Standard signals that carry information about neurodegeneration are signals from accelerometer measurements of limb tremors. The reason is the high sensitivity of commonly available sensors, mainly MEMS sensors, which ensure the recording of tremor even below the threshold of its visibility [12,13]. Unfortunately, the commonly attributed limb tremor to Parkinson's Disease does not occur in all patients. It is estimated that it affects about 75% of cases [14]. Hence, there is a search for other methods of data measurement and analysis [15]. Documented results of medical research indicating the impact of Parkinson's disease on the process of respiration, phonation and articulation open up a wide range of possibilities for the analysis of recorded signals using modern methods. These methods use both domain knowledge in the field of automatic recognition of speakers and speech [16,17] and convolutional networks to process images obtained as a result of time-frequency analysis [18]. The advent of modern graphics tablets has opened up the possibility of analyzing the dynamics and kinematics of writing towards micrography occurring in Parkinson's disease [19]. Similarly, the development of wearable and non-wearable sensor technologies, including inertial sensors, pressure and force sensors, ultrasonic sensors,

floor sensors and technologies based on image processing, including deep learning [20], allows the detection of various patterns of gait disturbances, e.g., gait freezing, reduced step length, problems with gait initiation, shuffling steps as well as balance and postural control [21].

2.2. Evaluation of Facial Bradykinesia in PD Patients

Studies on hypomimia in patients suffering from Parkinson's disease have been conducted for many years [6,22–24]. Their results show that, in general, PD patients show a reduced ability to spontaneously express their faces in everyday situations and in various experiments, including those designed to evoke emotional facial expressions. A fundamental work of recent years, pointing to the possibility of using hypomimia as a good biomarker in Parkinson's disease, is a paper published in 2021 by Maycas-Cepeda et al. [7]. In their research, including 75 PD patients, they investigated the relationship between hypomimia and the occurrence of motor and non-motor symptoms and achieved a high correlation with clinical evaluation performed using the UPDRS test and other tests. In particular, the Spearman correlation between hypomimia and clinical scales, including total and motor UPDRS, was 0.551 and 0.529, respectively, with a p -value below 0.01. On the basis of the obtained results, it was found that hypomimia may be a useful global marker of the overall severity of the disease, including cognitive decline.

Objectifying facial expressivity assessment of PD patients can be naturally achieved in a non-invasive way by using digital cameras together with image processing methods. One of the methods to measure changes in the area of a face caused by contractions of facial muscles is based on the Facial Action Coding System (FACS) developed by Ekman [25]. The human face has a number of muscles, the activation of which shows in the population certain regularities occurring in the expression of a specific type of emotion. Ekman's system assumes the existence of some elementary expressions of facial muscles, called unit activities AU. The AUs were used in some recent papers. Ali et al. [26], in 2021, analyzed videos of 61 PD patients and 543 healthy cases and measured the variances of selected facial action units—AU6 (cheek raiser), AU12 (lip corner puller) and AU4 (brow lowered) during three posed facial expressions, including smile, disgust and surprise. They achieved 95.6% accuracy in the recognition of PD patients with the use of an SVM classifier trained on the variances. In earlier work, Wu et al. [27] used a similar methodology in their preliminary study based on eleven AUs and included seven PD patients and eight control participants. The subjects were asked to produce neutral facial expressions and expressions that resemble amusement, sadness, anger, disgust, surprise and fear; however, only disgust was used in the analysis as its intensity was rated as the highest by the participants themselves while watching recorded videos.

In 2020, we also proposed the use of objective image processing techniques to assess face images recorded during a clinical trial [28]. The procedure of data acquisition used in facial masking examination assumed the registration of natural facial expression and posed facial expression uniform for all patients. Patients were asked to try to present a smile on their faces. The idea was to determine in the first place the key points of the face, which include points on the arches of the eyebrows and eyes, points of outline of the mouth, nose, etc. Next, it was necessary to determine the vector of characteristic features expressed by geometric relations between the points of the face. This vector changed with the emotion expressed on the face. An algorithm based on deep learning techniques (neural convolutional network) was used to determine the facial key points. In the development of the algorithm, 10,000 learning patterns were used [29]. The motivation for its use was the fact that the training data included images of faces with strongly different poses and representing variable facial expressions. The algorithm detects 68 points but requires a prior indication of the area of interest in the form of a border of the face image, which can be relatively easily obtained using the classic Viola–Jones algorithm. The proposed idea was significantly developed in 2021 and presented by Ge Su et al. [30]. They also used the 68 key points of the face to create geometric features. Using the SVM classifier, they obtained the

quality indicators of the diagnosis of hypomimia expressed by *precision*, *recall* and F_1 *score*, equal to 77.42%, 81.31% and 0.7931, respectively. They also used textural features and a fusion of geometric and textural features and achieved significantly better results.

3. Problem Statement

In the related papers presented above, the authors do not report the level of severity of disease symptoms among PD patients. In contrast, this study presents the results of research on the use of image processing methods to detect hypomimia using a research group of patients who were in the so-called switched ON phase. The ON phase is the action phase of drugs that reduce the symptoms of the disease. This poses a much greater challenge to the data processing methods being developed, the purpose of which is to automate the process of diagnosing Parkinson's disease.

The methods of machine learning proposed in this paper included, as a comparison, the approach presented in the paper [30], where using hand-crafted geometric features of images recorded in the visible light range in combination with the SVM classifier was applied. The advantage of geometric features is their easy interpretation; however, some domain knowledge is needed to indicate the relevant features for solving the recognition task. In the case of human faces, the selection of appropriate geometric features may be difficult due to the fact that they also contain individual information on personal identity.

Therefore, an approach based on the CNN convolutional network was also proposed, which is capable of automatically generating features in the learning process without the need for any step of pre-processing. In the case of the convolutional network, the research used facial images recorded in the visible light range, in the infrared range and images resulting from the fusion of both spectral ranges.

4. Materials and Methods

4.1. Image Acquisition

The data for analysis were obtained as part of clinical trials conducted by a medical team in the Department of Neurology at the Medical University of Warsaw in Poland after obtaining the consent of the Bioethics Committee. A total of 48 people took part in the study. There were 24 people in the PD group with Parkinson's disease. The health control (HC) group also included 24 people. The limited number of patients and healthy people who participated in data registration affected the way the convolutional network was used. A short discussion of the adopted strategy is given in Section 5.2. The basic characteristics of both groups are presented in Table 1.

Table 1. The basic characteristics of the research (PD) and control group (HC).

Volunteers	Male	Female	Total	Age
PD	8	16	24	61 ± 15
HC	19	5	24	43 ± 15

The result of the recordings carried out during the clinical trial using the system presented in Figure 1 were short films depicting the subjects with a natural facial expression marked further as A and facial expression with a smile marked further as B. Expression B was to reflect the subjective state of patients' satisfaction or happiness. In order to record a thermal image in the wavelength range from 7.5 to 13 μ m, a Flir A65 thermal imaging camera with a focal length of 25 mm was used. The camera provided a field of view of 25° × 20° and was equipped with a matrix with a resolution of 640 × 512. Visible light image recording was carried out using a Basler acA1440-220uc camera equipped with a lens providing a field of view with dimensions of 35° × 26°. The matrix of the Basler camera allowed the acquisition of a color image with a resolution of 1440 × 1080. The small size of both devices used for image recording made it possible to mount them on one photo tripod, which ensured a short distance of the optical axes of the lenses of both

cameras and allowed for the observation of patients from almost the same perspective in two spectral ranges.

Figure 1. The set used for image acquisition.

4.2. Selection of ROI Areas

Recordings were made simultaneously using both cameras. Static images were used in the calculations. Thermal imaging based on the recorded temperature distribution was obtained by mapping the temperature range from 20 to 30 degrees Celsius into the intensity range from 0 to 255 of a monochrome image with an 8-bit resolution. The recorded material was first subjected to pre-processing. The purpose of the pre-processing was to extract the fragments intended for further research, i.e., to determine the area of the face in both spectral ranges so as to obtain coverage not only of the dimensions of the cut fragments but above all the coverage of facial images. The results of the sample selection of ROI areas are depicted in Figure 2.

Figure 2. Corresponding areas of interest in (a) visible light image and (b) thermal image.

4.3. Methods Used to Quantify the Recognition Process

In order to reliably assess the proposed methods of recognizing patients with hypomimia in Parkinson's disease, the imaging material was divided into subsets used to train and evaluate the performance of classifiers. A commonly used cross-validation method was applied to assess the recognition process without the risk of obtaining too optimistic and unreliable assessments, especially when the dataset is not very numerous. The method consists in randomly dividing the entire dataset into N subsets of equal size. Then, a single subset is used to validate the model, while the remaining subsets are used in

training. The process is repeated N times in such a way that within each repetition, the validation and training sets are disjointed. The classification results obtained by means of validation subsets are then averaged. The N value is usually taken from 5 to 10, but there are no special rules governing the procedure in this respect [31]. In the described studies, in order to use test and training sets independent of the person, a single subset in cross-validation contained a total of 8 images corresponding to 4 people from each of the two classes. In this way, 6-fold cross-validation was obtained—data obtained from 40 people were used in training, and data from the remaining 8 people were used in the validation. At the same time, the requirement that the data of people from the validation group should not be included in the training group was met.

The results of the recognition, for comparison with the material presented in [30], were expressed by the *precision*, *recall* and F_1 score measures used there. These measures are taken from the concept of the error matrix [31]. The error matrix is a simple cross-tabulation of actual and recognized classes that allows classifier parameters to be easily calculated.

5. Results

5.1. Recognition Based on Geometric Features

The characteristics or key points in the facial image detected by the systems described in the literature are more or less related to the muscles responsible for facial expressions. Some of them, e.g., points describing the inner corners of the eyes, can theoretically be treated as points independent of the expressed emotion and used as stable reference points for those that lie in close proximity to the muscles responsible for facial expressions. In this way, the quantitative assessment of a person's ability to express emotions could theoretically be determined by the distance between properly selected characteristic points. This paper uses a system that detects 68 characteristic points, an example of which, for images with a natural facial expression A and with an expression B representing a smile, is presented in Figure 3.

Figure 3. Key points detected in images of the same person's face with different facial expressions: natural facial expression—A, smile—B.

Assuming that the line determined by the set of points detected along the nose is the axis of facial symmetry, it can be seen that some points, such as 40 and 43 or 2 and 16 with an exact frontal position of the face relative to the camera, should not change their position relative to this axis. They can thus become reference points for other points whose position changes under the influence of emotions. It should be noted, however, that due to natural head movements, faces at the time of recording may be at slightly different distances from the camera. They can also be tilted with different yaw, pitch and roll angles. In principle, frontal positioning of the face relative to the camera due to problems with maintaining a stable posture of sick people is very difficult. This makes the process of normalizing the distance between points on faces (e.g., by means of the distance between the inner

corners of the eyes) impossible to carry out. Therefore, in this paper, in contrast to [30], the expert selection of points that can potentially carry information about changing facial expressions was abandoned. Instead, an exploratory approach was used, which consisted in determining the distance between each pair of characteristic points in the images of both faces A and B, determining the characteristic features of pairs of images A and B from them, and then selecting the features in order to obtain the correct possible recognition. Relative changes in the distance between the corresponding points in images A and B were used as characteristic features. At 68 points, the number of pairs of points between which the distance can be determined is 2278. This is also the number of characteristic features for each pair of images on the basis of which the decision is made. The selection of features was made using the modified Fisher criterion J [32], in which instead of average and standard deviations of features in PD and HC populations, medians $m_{0.5}$ and interquartile ranges iqr were used:

$$J = \frac{|m_{0.5}(feature_{PD}) - m_{0.5}(feature_{HC})|}{iqr(feature_{PD}) + iqr(feature_{HC})} \quad (1)$$

A high J value indicates the high ability of a given feature to differentiate both PD and HC classes. In order to limit the number of features while maintaining as many correct diagnoses as possible, the features where the indicator exceeded the set threshold were selected in an experimental way. Figure 4 presents the relationship of the total accuracy of recognition ACC , understood as the number of all correct diagnoses related to the number of all people, i.e., people from the control group and the research group, as a function of the increasing threshold:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \cdot 100\%. \quad (2)$$

Figure 4. Accuracy of the recognition obtained in the test set as a function of the threshold value.

As a classifier, as in [30], an SVM neural network was used, which has mechanisms to prevent overfitting. For the optimal threshold of 0.08, which made it possible to achieve an accuracy of 83.3%, the number of features was 788. The error matrix for this optimal threshold is shown in Figure 5.

5.2. Recognition Using a Convolutional Network

As mentioned earlier, the motivation to use the convolutional network (CNN) in the diagnosis of hypomimia is the possibility of abandoning the methods of pre-processing, which are developed on the basis of expert knowledge about objects in the traditional machine learning approach. Instead, it automatically adjusts the filters of size and the number defined by the user. The natural strategy for building a CNN class solution is

to construct your own network structure from scratch, sufficiently developed to solve the task and train it using a large database of patterns. In the absence of such a base as in the problem of hypomimia recognition under this consideration, another alternative procedure called the transfer learning strategy is possible [20,33]. The idea is to use a network already trained to solve a completely different task as a base structure, which is then modified by fine-tuning operations. As part of the initial activities, it is necessary to download a network pre-trained using a large database of image patterns. Its first layers detect the so-called low-level features such as edges, colors and characteristic clusters of pixels, while the final layers generate features adequate to the task for which the network was prepared. In the second step, these last layers should be removed and replaced with a network structure not yet trained but adequate to the task being solved [10,34]. In the case of Parkinson's disease diagnosis, this structure must be selected so that it distinguishes only two categories, i.e., "sick" and "healthy". The third step is to train such a modified network using a new database. Step four is to examine how the network works in recognition mode.

geometrical features + SVM

recognized class	sick	20 41.7%	4 8.3%	83.3% 16.7%
	healthy	4 8.3%	20 41.7%	83.3% 16.7%
		83.3% 16.7%	83.3% 16.7%	83.3% 16.7%
		sick	healthy	
		actual class		

Figure 5. Error matrix showing the results of hypomimia recognition based on geometric features of facial images.

Networks trained on a large database include such net solutions as AlexNet, GoogLeNet, VGG-16 and VGG-19. However, application of them is not always easy due to the extensive layer structures that require the use of strong GPU processors to carry out the learning process, which can take a very long time, even in the transfer strategy of learning. For this reason, in this work, the AlexNet [35] network was used. It contains eight layers whose weights are subject to modification in the training process: five convolutional layers and three layers of full connection. For comparison, the VGG-19 network has 16 convolutional layers and 3 fully connected layers. Originally, the AlexNet network was adapted to distinguish a thousand classes of objects, while the problem under consideration in this paper requires the distinction of only two classes: a sick person and a healthy person.

The main problem related to the use of commonly available networks such as AlexNet is the adaptation of the input data to the requirements of the network. The available image data are color images with facial expressions A and B and corresponding monochrome images obtained from a thermal imaging camera based on the temperature distribution on the face surface. The AlexNet network is fed with color images made of three channels R, G and B, which in general can be created using any monochrome image. By having monochrome versions of color images A and B, marked later as *monVIS_A* and *monVIS_B*, as well as corresponding monochrome thermal images, marked as *monIR_A* and *monIR_B*, it is possible to use their combination to fill the R, G and B channels of a certain resulting image. Such an image will thus contain complex information about the object being studied, particularly information concerning the ability of faces to express emotions. A list of such monochrome components for the selected person is presented in Table 2.

Table 2. Proposed components for filling RGB channels on the example of a selected person.

<i>monVIS_A</i>	<i>monVIS_B</i>	<i>monIR_A</i>	<i>monIR_B</i>

As a result, in the research using the AlexNet network, color images obtained as part of direct recordings, marked as *rgbVIS_A* and *rgbVIS_B*, were mandatory used as input images. Besides, six different types of images resulting from the fusion of the components presented in Table 1, including the fusion of visible light and infrared images, were also used. They were marked from *fusion 1* to *fusion 6*. Fusions from 1 to 3 were dominated by images acquired in visible light, and the infrared images dominated in fusions from 4 to 6. Examples of imageries showing the proposed image combinations are depicted in Table 3, together with formulas for filling RGB channels.

Table 3. Examples of RGB class images used as AlexNet input.

<i>rgbVIS_A</i>	fusion 1 R: <i>monVIS_A</i> G: <i>monVIS_B</i> B: (<i>monVIS_A</i> + <i>monVIS_B</i>)/2	fusion 3 R: <i>monVIS_A</i> G: <i>monVIS_B</i> B: <i>monIR_A</i>	fusion 5 R: <i>monVIS_A</i> G: <i>monVIS_B</i> B: (<i>monIR_A</i> + <i>monIR_B</i>)/2

ACC = 89.6%	ACC = 93.8%	ACC = 93.8%	ACC = 93.8%
<i>rgbVIS_B</i>	fusion 2 R: <i>monIR_A</i> G: <i>monIR_B</i> B: (<i>monIR_A</i> + <i>monIR_B</i>)/2	fusion 4 R: <i>monIR_A</i> G: <i>monIR_B</i> B: <i>monVIS_A</i>	fusion 6 R: <i>monIR_A</i> G: <i>monIR_B</i> B: (<i>monVIS_A</i> + <i>monVIS_B</i>)/2

ACC = 89.6%	ACC = 85.4%	ACC = 85.4%	ACC = 85.4%

In order to adapt the AlexNet network, in the first step, its last three layers, i.e., the fully connected layer, SoftMax and the classification layer, were removed and replaced with new ones. The new layers, together with the remaining layers, were then modified in the process of training the network using the cross-validation scheme. We also used the common practice of reducing the learning rate because we did not want to distort the weights of the pre-trained network too quickly and too much. We set the learning factor

at 10^{-4} . However, the training process appeared to be long and provided the test results expressed by the Formula (2) for ACC at a level of 83.3%, i.e., at the level of recognition accuracy obtained when using geometric features.

Therefore, in the second step, research was performed for a network of decreasing size. The best result was obtained for a network containing only one convolutional layer. The final structure of the network is presented in Table 4.

Table 4. Final structure of the CNN network used.

Layer Number	Layer Type	Description
1	'data' Image Input	$227 \times 227 \times 3$ images with 'zerocenter' normalization
2	'conv1' Convolution	96 $11 \times 11 \times 3$ convolutions with stride [4 4] and padding [0 0 0 0]
3	'relu1' ReLU	ReLU
4	'pool1' Max Pooling	3×3 max pooling with stride [2 2] and padding [0 0 0 0]
5	Fully Connected	2 fully connected layer
6	Softmax	softmax
7	Classification Output	crossentropyex

Perfectly accurate results were obtained in the training process for each type of image. The curves representing this phase are depicted in Figure 6. During training, we used data augmentation based on a random rotation. Source images were rotated in the range of $\pm 10^\circ$, but this operation did not improve the results. The results of tests varied, depending on the type of input image. The quantitative values represented by the measure ACC expressed by Formula (2) are included in Table 2 under each type of image.

Figure 6. Basic characteristics of training the CNN convolutional network—the accuracy and loss curves.

The best results in terms of ACC were achieved for types of imagery called *fusion 1*, *fusion 3* and *fusion 5*. At the same time, the ACC were equal for these cases. Error matrices, slightly different for each of them, are depicted in Figure 7.

Figure 7. Error matrices obtained when recognizing hypomimia using CNN and fusion images: (a) *fusion 1*, (b) *fusion 3*, (c) *fusion 5*.

6. Discussion

The use of ACC accuracy measure allows us to conclude that regardless of the type of assumed input image, the approach based on the use of a convolutional network gives better results than the approach based on geometric features. At the same time, it was observed that images created from data acquired by the thermal camera give worse accuracy than data in which visible light images were dominant. These were even worse than the results obtained using raw RGB images of type A or B. Perhaps the processes of accumulation of pathological alpha-synuclein described in the literature do not occur significantly in the skin on the human face [8,9]. The values of the quality measures for the selected processing methods are summarized in Table 5.

Table 5. The performance of different frameworks.

Framework	Precision [%]	Recall [%]	F ₁ Score [-]
geometric features + SVM	83.3	83.3	0.833 ± 0.022
geometric features + SVM ¹	77.4 ¹	81.3 ¹	0.793 ± 0.030 ¹
<i>fusion 1</i> + CNN	100.0	88.9	0.941 ± 0.021
<i>fusion 3</i> + CNN	95.8	92.0	0.939 ± 0.037
<i>fusion 5</i> + CNN	91.7	95.7	0.937 ± 0.038
fusion features + SVM ¹	99.94 ¹	100.0 ¹	0.999 ± 0.000 ¹

¹ Results are presented in [30] without any information concerning the ON/OFF state of patients.

In the context of previous studies presented in [30], the results obtained in this work seem to be better as far as the use of geometric features and the SVM classifier is concerned. At the same time, the results of the use of the CNN network are characterized by slightly lower quality indicators than the best presented in [30], which were the result of the use of a fusion of hand-crafted features based on geometric and textural approaches. However, the cited results cannot be fully comparable to ours due to the fact that all of our patients participating in the studies were intentionally tested while being in the ON phase. The authors of the paper [30] did not mention this. The ON phase is characterized by reduced symptoms of the disease and imposes increased requirements on processing algorithms. It is impossible to state unequivocally which of the methods will be characterized in practice by better generalization abilities without performing additional comparative studies.

7. Conclusions

The presented material presents the possibilities of using various methods of image processing as methods supporting the diagnosis of patients with Parkinson's disease based on the detection of hypomimia. The results extend the achievements discussed in the most

comparable article [30] devoted to the use of image processing methods in this task. The novelty of this article is using a convolutional network fed with complex data created on the basis of compositions of images containing natural facial expressions and facial expressions with a smile. Although we obtained lower quality indicators expressed by the measure F_1 score: 0.941 using CNN vs. 0.999 using the fusion of geometric and textural features, it should be emphasized that our results are the result of research in a group of patients in the ON phase. It seems, therefore, that the results presented in this article are of greater value for the development of new diagnostic methods, different from those used during medical history, which uses assessment methods validated in the OFF phase, i.e., in the phase with clearly visible symptoms.

At the same time, it should be stated that the presented methods are attractive from the point of view of remote medical examination due to the fact that they do not require personal contact between the patient and the doctor. In future studies, we plan to develop regression models associated with the UPDRS score, which can be used to assess the severity of the disease.

Author Contributions: Conceptualization, A.P.-C. and J.J.; methodology, J.J. and A.P.-C.; software, K.B.; validation, A.K.-P.; formal analysis, J.J. and K.B.; investigation, K.B.; resources, A.P.-C. and M.N.; data curation, K.B.; writing—original draft preparation, J.J. and K.B.; writing—review and editing, J.J., A.P.-C., M.N. and A.K.-P.; visualization, K.B.; supervision, A.K.-P. and J.J.; project administration, J.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Polish Ministry of National Defense for the implementation of basic research within the research grant No. GBMON/13-996/2018/WAT “Basic research in the field of sensor technology using innovative data processing methods”.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Pham, H.N.; Do, T.T.T.; Chan, K.Y.J.; Sen, G.; Han, A.Y.K.; Lim, P.; Cheng, T.S.L.; Nguyen, Q.H.; Nguyen, B.P.; Chua, M.C.H. Multimodal Detection of Parkinson Disease based on Vocal and Improved Spiral Test. In Proceedings of the International Conference on System Science and Engineering, Dong Hoi City, Vietnam, 19–21 July 2019; pp. 279–284. [CrossRef]
2. Gawel, M.; Potulska-Chromik, A. Neurodegenerative diseases: Alzheimer’s and Parkinson’s disease. *Postępy Nauk. Med.* **2015**, *28*, 468–476.
3. Maśliński, S.; Ryżewski, J. *Pathophysiology*, 4th ed.; PZWL: Warszawa, Poland, 1992.
4. Obeso, J.A. The Unified Parkinson’s Disease Rating Scale (UPDRS): Status and recommendations. *Mov. Disord.* **2003**, *18*, 738–750.
5. International Parkinson and Movement Disorder Society: “MDS Rating Scales”. Available online: <https://www.movementdisorders.org/MDS/Education/Rating-Scales.html> (accessed on 20 October 2021).
6. Gunnery, S.D.; Habermann, B.; Saint-Hilaire, M.; Thomas, C.A.; Tickle-Degnen, L. The Relationship between the Experience of Hypomimia and Social Wellbeing in People with Parkinson’s Disease and their Care Partners. *J. Parkinson’s Dis.* **2016**, *6*, 625–630. [CrossRef]
7. Maycas-Cepeda, T.; López-Ruiz, P.; Feliz-Feliz, C.; Gómez-Vicente, L.; García-Cobos, R.; Arroyo, R.; García-Ruiz, P.J. Hypomimia in Parkinson’s Disease: What Is It Telling Us? *Front. Neurol.* **2021**, *11*, 1775. [CrossRef]
8. Nolano, M.; Provitera, V.; Manganelli, F.; Iodice, R.; Stancanelli, A.; Caporaso, G.; Saltalamacchia, A.; Califano, F.; Lanzillo, B.; Picillo, M.; et al. Loss of cutaneous large and small fibers in naive and l-dopa-treated PD patients. *Neurology* **2017**, *89*, 776–784. [CrossRef]
9. Purup, M.M.; Knudsen, K.; Karlsson, P.; Terkelsen, A.J.; Borghammer, P. Skin Temperature in Parkinson’s Disease Measured by Infrared Thermography. *Parkinson’s Dis.* **2020**, *2020*, 2349469. [CrossRef]
10. Khan, M.A.; Imran, A.; Majed, A.; Damaševičius, R.; Scherer, R.; Rehman, A.; Bukhari, S. Multimodal Brain Tumor Classification Using Deep Learning and Robust Feature Selection: A Machine Learning Application for Radiologists. *Diagnostics* **2020**, *10*, 565. [CrossRef]
11. Khan, M.A.; Akram, T.; Sharif, M.; Kadry, S.; Nam, Y. Computer decision support system for skin cancer localization and classification. *Comput. Mater. Contin.* **2021**, *68*, 1041–1064. [CrossRef]
12. Pierleoni, P. A Smart Inertial System for 24h Monitoring and Classification of Tremor and Freezing of Gait in Parkinson’s Disease. *IEEE Sens. J.* **2019**, *19*, 11612–11623. [CrossRef]
13. Kumar, Y.D.; Prasad, A.M. MEMS accelerometer system for tremor analysis. *Int. J. Adv. Eng. Glob. Technol.* **2014**, *2*, 685–693.

14. Zach, H.; Dirks, M.; Bloem, B.R.; Helmich, R.C. The Clinical Evaluation of Parkinson's Tremor. *J. Parkinson's Dis.* **2015**, *5*, 471–474. [\[CrossRef\]](#)
15. Chmielińska, J.; Białek, K.; Potulska-Chromik, A.; Jakubowski, J.; Majda-Zdancewicz, E.; Nojszewska, M.; Kostera-Pruszczyk, A.; Dobrowolski, A. Multimodal data acquisition set for objective assessment of Parkinson's disease. In Proceedings of the SPIE 1142, Radioelectronic Systems Conference 2019, Jachranka, Poland, 11 February 2020; p. 114420G. [\[CrossRef\]](#)
16. Gunduz, H. Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets. *IEEE Access* **2019**, *7*, 115540–115551. [\[CrossRef\]](#)
17. Das, B.; Daoudi, K.; Klempir, J.; Rusz, J. Towards disease-specific speech markers for differential diagnosis in parkinsonism. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing 2019, Brighton, UK, 12–17 May 2019; pp. 5846–5850.
18. Majda-Zdancewicz, E.; Potulska-Chromik, A.; Jakubowski, J.; Nojszewska, M.; Kostera-Pruszczyk, A. Deep Learning vs. Feature Engineering in the Assessment of Voice Signals for Diagnosis in Parkinson's Disease. *Bull. Pol. Acad. Sci. Tech. Sci.* **2021**, *69*, e137347. [\[CrossRef\]](#)
19. Drotar, P.; Mekyska, J.; Rektorová, I.; Masarová, L.; Smekal, Z.; Faundez-Zanuy, M. Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. *Artif. Intell. Med.* **2016**, *67*, 39–46. [\[CrossRef\]](#)
20. Khan, M.A.; Kadry, S.; Parwekar, P.; Damaševičius, R.; Mehmood, A.; Khan, J.A.; Naqvi, S.R. Human gait analysis for osteoarthritis prediction: A framework of deep learning and kernel extreme learning machine. *Complex. Intell. Syst.* **2021**, 1–19. [\[CrossRef\]](#)
21. Di Biase, L.; Di Santo, A.; Caminiti, M.L.; De Liso, A.; Shah, S.A.; Ricci, L.; Di Lazzaro, V. Gait Analysis in Parkinson's Disease: An Overview of the Most Accurate Markers for Diagnosis and Symptoms Monitoring. *Sensors* **2020**, *20*, 3529. [\[CrossRef\]](#)
22. Aarsland, D.; Ballard, C.; McKeith, I.; Perry, R.H.; Larsen, J.P. Comparison of extrapyramidal signs in dementia with Lewy bodies and Parkinson's disease. *J. Neuropsychiatry Clin. Neurosci.* **2001**, *13*, 374–379. [\[CrossRef\]](#)
23. Simons, G.; Pasqualini, M.C.S.; Reddy, V.; Wood, J. Emotional and nonemotional facial expressions in people with Parkinson's disease. *J. Int. Neuropsychol. Soc.* **2004**, *10*, 521–535. [\[CrossRef\]](#)
24. Hunker, C.J.; Abbs, J.H.; Barlow, S.M. The relationship between parkinsonian rigidity and hypokinesia in the orofacial system: A quantitative analysis. *Neurology* **1982**, *32*, 749–754. [\[CrossRef\]](#)
25. Ekman, P.; Freisen, W.V.; Ancoli, S. Facial signs of emotional experience. *J. Personal. Soc. Psychol.* **1980**, *39*, 1125–1134. [\[CrossRef\]](#)
26. Ali, M.R.; Myers, T.; Wagner, E.; Ratnu, H.; Dorsey, E.R.; Hoque, E. Facial expressions can detect Parkinson's disease: Preliminary evidence from videos collected online. *npj Digit. Med.* **2021**, *4*, 1–4. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Wu, P.; González, I.; Patsis, G.; Jiang, N.; Sahli, H.; Kerckhofs, E.; Vandekerckhove, M. Objectifying facial expressivity assessment of Parkinson's patients: Preliminary study. *Comput. Math. Methods Med.* **2014**, *2014*, 427826. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Białek, K.; Jakubowski, J.; Potulska-Chromik, A.; Chmielińska, J.; Majda-Zdancewicz, E.; Nojszewska, M.; Kostera-Pruszczyk, A.; Dobrowolski, A. Selected problems of image data preprocessing used to perform examination in Parkinson's disease. In Proceedings of the SPIE 1142, Radioelectronic Systems Conference 2019, Jachranka, Poland, 11 February 2020; p. 114420G. [\[CrossRef\]](#)
29. Zhang, Z.; Luo, P.; Loy, C. Facial Landmark Detection by Deep Multi-task Learning. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 94–108.
30. Su, G.; Lin, B.; Yin, J.; Luo, W.; Xu, R.; Xu, J.; Dong, K. Detection of hypomimia in patients with Parkinson's disease via smile videos. *Ann. Transl. Med.* **2021**, *9*, 1307. [\[CrossRef\]](#)
31. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*, 1st ed.; Springer: New York, NY, USA, 2013.
32. Hegde, S.; Achary, K.K.; Shetty, S. Feature selection using Fisher's ratio technique for automatic speech recognition. *Int. J. Cybern. Inform.* **2015**, *4*, 45–52. [\[CrossRef\]](#)
33. Shao, L.; Zhu, F.; Li, X. Transfer Learning for Visual Categorization: A Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 1019–1034. [\[CrossRef\]](#)
34. Ramdan, A.; Heryana, A.; Arisal, A.; Kusumo, R.B.S.; Pardede, H.F. Transfer Learning and Fine-Tuning for Deep Learning-Based Tea Diseases Detection on Small Datasets. In Proceedings of the International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET), Tangerang, Indonesia, 18–20 November 2020; pp. 206–211. [\[CrossRef\]](#)
35. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [\[CrossRef\]](#)