IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 20, 2019, accepted January 6, 2020, date of publication January 21, 2020, date of current version January 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968367

TestGraphia, a Software System for the Early Diagnosis of Dysgraphia

GIOVANNI DIMAURO[1], VITOANTONIO BEVILACQUA[2], LUCIO COLIZZI[1], AND DAVIDE DI PIERRO[1]

[1]Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro, 70125 Bari, Italy
[2]Dipartimento di Ingegneria Elettrica e dell'Informazione, Politecnico di Bari, 70125 Bari, Italy

Corresponding author: Giovanni Dimauro (giovanni.dimauro@uniba.it)

ABSTRACT Dysgraphia, which is known as a writing disorder, is a specific disorder of writing regarding the reproduction of alphabetical and numerical signs. Dysgraphia may be related to dyspraxia, which is secondary to incomplete lateralization and characterized by a difficulty to reproduce alphabetical and numerical signs. Since the causes of dysgraphia are unknown, the rapid detection of symptoms is very important. In academic and clinical uses, the most common tool for detecting dysgraphia is an evaluation of the quality of writing on paper sheets. A writing analysis is based on rules for scoring the writing quality. In this paper, we discuss TestGraphia, which is a software system that can support doctors in making diagnoses and monitoring patients with dysgraphia in an objective manner. The system is based on known document analysis algorithms and modified or specially designed algorithms. Based on this software, a forms analysis requires considerably less time than that needed by traditional methods, enabling large screening activities and reducing time and cost. Potential dynamic changes in dysgraphia screening can be assessed by monitoring the quality of writing in a non-invasive way with reduced costs, both in the laboratory and the patient's home, and the appropriate frequency. In the system that we will describe, the mean time to execute a diagnosis is nearly ten times faster with trustworthy results.

INDEX TERMS Dysgraphia, document analysis system, BHK test, handwriting analysis.

I. INTRODUCTION

Dysgraphia is a specific learning disorder that causes difficulties in reproducing both alphabetical and numerical signs; it concerns not only graphics but also indirect spelling and syntactic rules due to frequent challenges with rereading and self-correction. Children with signs of dysgraphia write irregularly; their hands flow with difficulty on a writing surface; the handle of their writing medium is often incorrect; the positions of their bodies is inadequate in most cases; their elbows are not placed on the table; and their torsos are excessively inclined [1]. Handwriting difficulty or dysgraphia was defined by Hamstra-Bletz and Blote as a disturbance or difficulty in the production of written language that is related to the mechanics of writing. This disorder has also been referred to as a specific learning disability. The problem is manifested in the inadequate performance of handwriting among children who have at least average intelligence and who have not been identified as having any distinct neurological problems. If their handwriting is very slow, children may forget the ideas and plans held in their memory before they succeed in transferring them to paper [2].

A distinction between "surface" and "deep" dysgraphia, which is made based on linguistic errors, has been used to support the argument that two main cognitive systems interact to enable proper translation of mental language to written language. The first cognitive system is considered to be a phoneme-to-grapheme translation system. In deep dysgraphia, the phonologic system is inoperative, whereas in surface dysgraphia, the lexical system is disrupted [3]. Comprehensive neuropsychological and electroencephalo-graphic examination of children with dysgraphia and dyslexia demonstrated local electroencephalo-graphy (EEG) anomalies in various zones of the cortex in both hemispheres, mainly in the posterotemporal and/or anterior regions of the left hemisphere and the posterotemporal regions of the right hemisphere. The character of speech disturbances considerably depended on the localization of the baseline EEG anomalies: disturbances of motor components of writing prevailed when these anomalies were localized in the anterior parts of the left

The associate editor coordinating the review of this manuscript and approving it for publication was Omar Sultan Al-Kadi.

hemisphere, while phonological and morphological speech disturbances were noted in children with local anomalies in the temporal regions of both hemispheres and in posterior association regions, mainly in the right hemisphere [4]. This disease can also cause other learning problems, such as dyslexia, attention deficit and dyspraxia. Developmental dysphasia, dyslexia and dysgraphia are more commonly associated in affected children than children who are not affected. Different members of the same family may show rather different manifestations of these disorders; for example, a dysphasic child may have a family history of specific developmental dyslexia and dysgraphia but have no dysphasia [5]. Children with dysgraphia also often have social problems since they feel less capable than other children [6].

The treatment can vary and includes motor exercises to coordinate the hand with the eye. Doctors recommend using a computer instead of paper to learn. The use of computers to facilitate intensive repetitive practice proved to be successful, both in terms of improvements on assessments and evidence of functional benefits. The use of a dictionary to support the strategy and an adaptive word processor to promote functional carryover is described. The role of the computer in therapy is discussed as a tool to facilitate repetitive practice of therapy and encourage independent use of the strategy embodied in therapy [7].

Sounds can be used to inform about the correctness of an ongoing movement without directly interfering with visual and proprioceptive feedback [8]. Evidence from recent studies suggests that writing and speaking may be an aspect of cognition that is capable of identifying impairments specific to patients with Parkinson and Alzheimer disease [9]–[11] or spatial dysgraphia in patients with right hemisphere lesions [12].

Danna et al. collected some physical features of the written capabilities of a child, such as: velocity difference in signal-to-noise velocity peaks and jerk movements. Starting with pen movements, some information about sound is collected. This method translates analogical movement into sound to detect anomalies [13]. Sound and technologies are also used to support aphasic persons. In one experiment, a 63-year-old man with fluent aphasia and severe acquired dysgraphia and dyslexia had poor social participation and was unable to work. Treatment consisted of 16 one-hour sessions. The man was trained to use Dragon Naturally Speaking VRS to assist writing and Read+ WriteGold text-to-speech software to assist with reading the development of the computer skills required to use e-mail. Outcome measures evaluated writing efficiency and communicative effectiveness, the functional impact of intervention, and changes in participation [14]. Augmentative and Alternative Communication (AAC) has primarily been utilized for motor speech deficits or as an aid for communicating basic needs in the acute stages of aphasia rehabilitation [15]. Other important studies correlate handwriting movements to the activation of the cerebral regions associated with the production of writing movement. The kinematic parameters of handwriting movements were directly correlated with functional magnetic resonance imaging examination as applied in other medical fields [16]–[18]. The findings of the previous study seem to be promising for evaluating the handwriting movement deficits and potential alterations in the neural activity in individuals with handwriting difficulties.

In literature there are also different automatic systems that aim to diagnose dysgraphia. In [24], [25] and [26] detailed diagnosis processes are descripted. The authors determine many parameters using Machine Learning and Information Theory's techniques. The previous systems detect dysgraphia with different methods than BHK, which is now one of the most widely used. In addition, other systems are certainly effective, but use electronic devices and therefore are less advantageous in carrying out screening than systems based on the analysis of manuscript documents. Further considerations will be discussed later in the paper.

In the following sections, we define the method applied in this study and describe how our system works.

II. BHK EVALUATION PROTOCOL

The Beknopte Beoordelingsmethode voor Kinderhandschriften (BHK) method was introduced in the Netherlands in 1987 at the Department of Developmental Psychology of the University of Leiden [19]. This method is the most frequently employed method for evaluating the quality of writing in academic and clinical applications. Thirteen parameters are scored to detect the features of a graphical act, which enables the detection of "dysgraphia". The estimation process that permits the detection and quantification of the principal features of poor writing is based on a standardized text to be copied and a few tools: white paper without lines, A4 format, physical support, a black pen and a timer. The position of the child who is being evaluated needs to be appropriate for the writing surface. The Italian standardized text is a re-adaptation of the Dutch text; the first five sentences are *leo e lo zio, sono al porto, mangiano un gelato, con loro ci sono,* and *mia e rina.* The child is instructed to copy the printed text in italics on white paper in five minutes and at least the first five lines. If the child fails to finish the first five sentences in five minutes, he will be granted additional time. When the final rating (refer to subsequent sections) shows a standard deviation (SD) below the average (-1.5 SD), then the writing is considered hard to read and "dysgraphic". The case of a slightly lower performance can be considered poor quality writing but not "dysgraphic" [13]. This method can be influenced by the interpretation and experience of the specialist, albeit with a minor impact on the final result. For this reason, in [13] an agreement value among four specialists was estimated.

The most significant task is to determine the scores of the following thirteen handwriting features that enable a diagnosis: (1) writing size, (2) non-aligned left margin, (3) skewed writing, (4) insufficient space between two words, (5) sharp angles, (6) broken links between two letters, (7) collision between two letters, (8) irregular size of letters,

(9) inconsistent height between letters with extension and letters without extension, (10) atypical letters, (11) ambiguous letters, (12) traced letters, and (13) unstable track. All but one feature can be scored between 0 and 5, while feature 9 has a maximum value of 4. Note that certain features are closely geometry-based, while other features require a doctor's interpretation and some features at this phase can be automized.

A. WRITING SIZE

The writing size is related to the body heights of the characters. We can discern various levels depending on the average height, and the feature score will depend on a child's class; the scores are listed in Table 1 as shown in [13]. Figure 1 reports a writing style in which the font size is larger than normal.

TABLE 1. Writing size reference (©2011) Erickson.

Primary School Class	<3 mm	4 mm	5 mm	6 mm	7 mm	8 mm	9 mm
First (5-6 years old)	0	0	1	2	3	4	5
Second	0	1	2	3	4	5	5
third	0	1	2	3	4	5	5
fourth	0	1	2	3	4	5	5
fifth	0	1	2	3	4	5	5

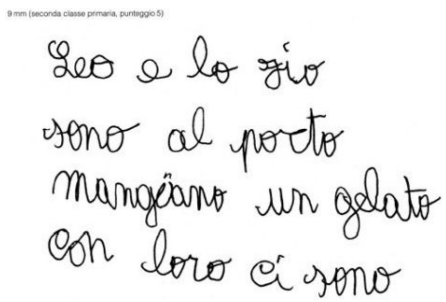

FIGURE 1. Size of characters is above the norm (©*2011 Erickson*).

In Figure 3 the average height of characters is ≈10mm.

B. NON-ALIGNED LEFT MARGIN

The left margin should be vertically aligned but sometimes can be tilted to the right. The rating is dependent on the entire written text. In the case of an irregular margin, the feature score is 0. In the case of its inclination to the right, the appropriate score is determined by means of a transparent matrix, as shown in Figure 2. An example of an inclined margin is shown in Figure 3.

C. SKEWED WRITING

The writing may be undulated, as shown in Figure 4. In this example, a word goes up or down. The value is determined by considering, if some characters are too far under or over the baseline for each line. The baseline is designed by connecting the first and last character of a line with the lowest points. If an irregularity is detected in a line, a score of 1 will be assigned; otherwise, a score of 0 will be assigned. The feature score will be the sum of the line values.

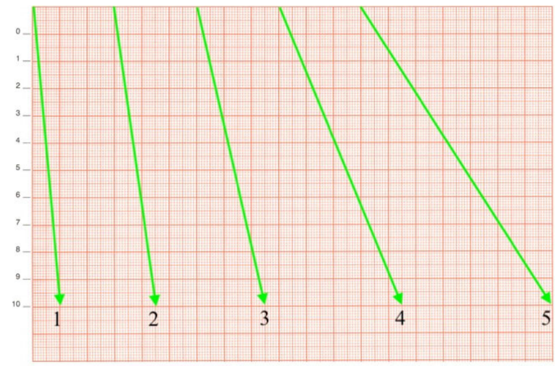

FIGURE 2. Transparent matrix (©*2011 Erickson*).

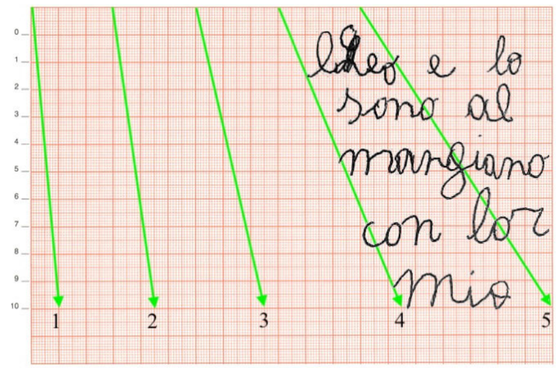

FIGURE 3. Non-aligned writing.

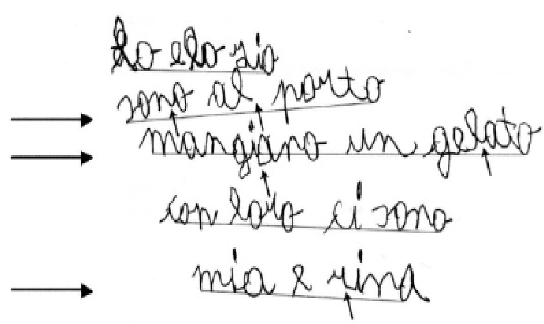

FIGURE 4. Skewed writing (©2011 Erickson).

D. INSUFFICIENT SPACE BETWEEN WORDS

The gap between two words is determined to be inadequate (Figure 5) if it is smaller than the width of a reference character, typically the letter "o". For every line, a value of 1 is assigned if the distance between two words is insufficient; otherwise, a value of 0 is assigned. The feature score will be calculated as the sum of the line values.

E. SHARP ANGLES

Sharp angles are stretched horizontal links or sharp angles that are present rather than curvy ones (Figure 6). For every

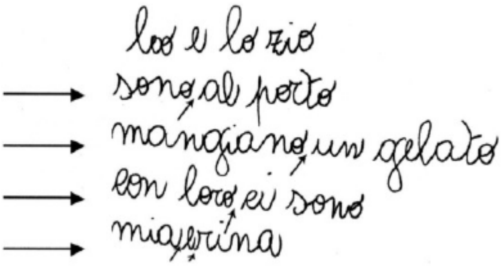

FIGURE 5. Text with insufficient spaces between two words (©*2011 Erickson*).

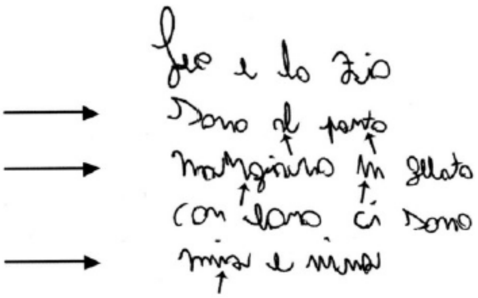

FIGURE 6. Text with sharp angles (©*2011 Erickson*).

line, we assign a value of 1 if sharp angles exist; otherwise, we assign a value of 0. The total score will be the sum of the line values.

F. BROKEN LINKS BETWEEN LETTERS

Broken links are generated when the pen motion is stopped or when the pen is shifted from paper (Figure 7), which may be caused by an unexpected change in the writing path or a missing binding of letters of a word. For each line, a value of 1 is assigned if interrupted links exist; otherwise, a value of 0 is assigned. The score will be the sum of the line values.

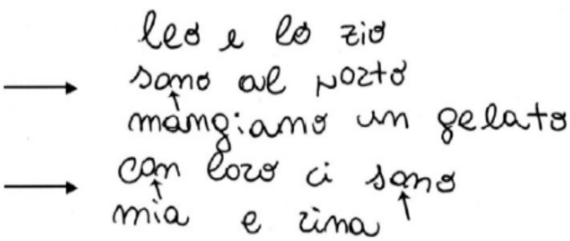

FIGURE 7. Text with broken links between words (©*2011 Erickson*).

G. COLLISIONS BETWEEN TWO LETTERS

Two characters collide if they are very near each other to create an overlaying area (Figure 8). Each line will be assigned a value of 1 if a collision occurs between two characters; otherwise, a value of 0 is assigned. The feature score will be the sum of the line values.

H. IRREGULAR SIZE OF LETTERS

When the letter body is not evenly sized (Figure 9), irregularity can occur within two characters of the same word or two characters within a row. For each row, the size is

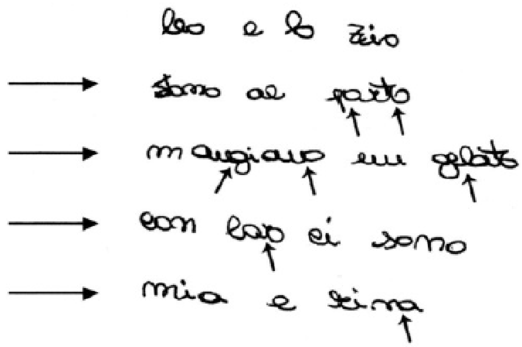

FIGURE 8. Text with collisions between two letters (©*2011 Erickson*).

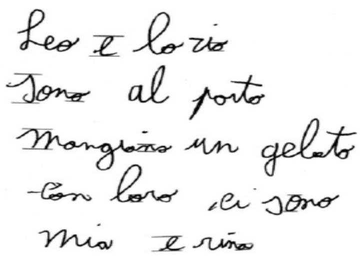

FIGURE 9. Text with irregular size of letters (©2011 Erickson).

considered the height of the highest letter and the height of the lowest, without elongations. Based on the dimension of the lowest character, the highest character should not exceed a maximal size. For each line, if the size of the highest character overcomes this limit, a value of 1 is assigned; otherwise, a value of 0 is assigned. The feature score will be the sum of the line values. Table 2 shows the correct size of the letters.

TABLE 2. Letters' size reference. (©*2011 Erickson).*

body height of the letters (mm)			
smallest	highest	smallest	highest
1,5	2,5	5,5	7,5
2	3	6	8
2,5	4	6,5	8,5
3	4,5	7	9,5
3,5	5	7,5	10
4	6	8	11
4,5	6,5	9	12
5	7		

I. INCONSISTENT HEIGHT BETWEEN LETTERS WITH AND WITHOUT AN EXTENSION

In this case, a slight variation in size exists between characters with an extension and characters without an extension (Figure 10). A value of 1 will be assigned to each line if the divergence is not very meaningful; otherwise, a value of 0 is assigned. The feature score will be obtained from the sum of the values of the rows.

J. ATYPICAL LETTERS

A character is defined "atypical" if it is considerably different from a common style (Figure 11). For each line, a value of 1 will be assigned if atypical letters exist; otherwise, a value

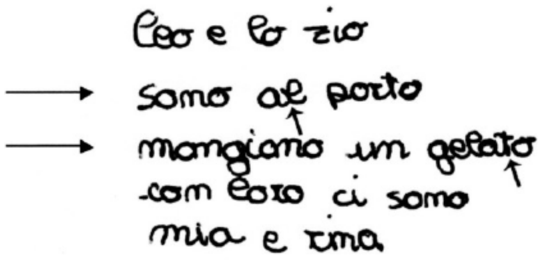

FIGURE 10. Inconsistent height between letters with an extension and letters without an extension (©*2011 Erickson*).

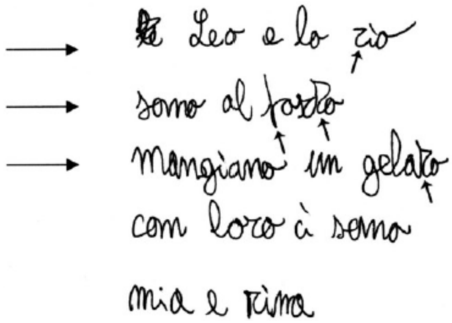

FIGURE 11. Text that contains atypical letters (©*2011 Erickson*).

of 0 is assigned. The feature score will be the sum of the values of the rows.

K. AMBIGUOUS LETTERS

Ambiguous letters are letters that are included in the reference alphabet but can cause problems in the interpretation of the characters (Figure 12). For each line, a value of 1 will be assigned if ambiguous letters exist; otherwise, a value of 0 will be assigned. The feature score will be the sum of the values of the rows.

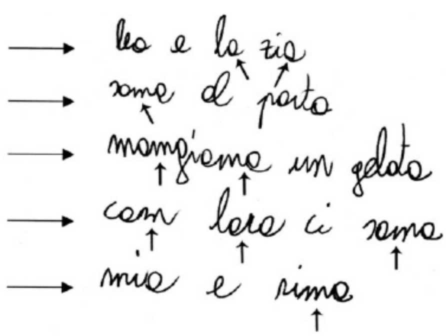

FIGURE 12. Some ambiguous letters are indicated (©*2011 Erickson*).

L. TRACED LETTERS

A character is defined as "traced" if it has been partly or completely redrafted to fit its form (Figure 13). For each line, a value of 1 will be assigned if some letters drawn; otherwise a value of 0 will be assigned. The feature score will be the sum of the values of each row.

FIGURE 13. Text with traced letters (©*2011 Erickson*).

M. UNSTABLE TRACK

An unstable track occurs when the writing is insecure, wavering or wobbly (Figure 14). For each line, a value of 1 is assigned if an unstable track is located on the row; otherwise, a value of 0 is assigned. The feature score will be the sum of the row values.

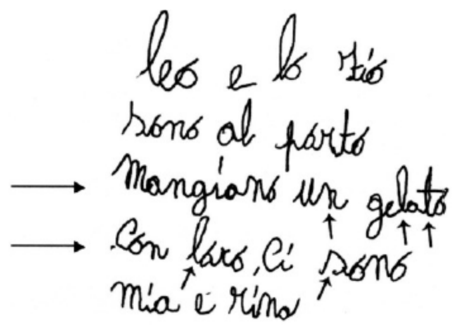

FIGURE 14. This text has some unstable tracks (©*2011 Erickson*).

III. SYSTEM AND METHODOLOGY

The purpose of TestGraphia is to automatically evaluate certain features and easily set the remaining features to simplify the diagnosis. The doctor must check only a few features that need to be interpreted. At the conclusion of the evaluation procedure, a final report that describes the findings of the system, such as the anomalies, features score, and total score.

A one-line header must be added to the default text, which contains the following characters: "a a o o m m". The purpose of this line is two-fold: to find the first feature score, the average value of the size of the characters must be computed without stretching them; second, for the feature 4, the average width of the character "o" must be determined by averaging the widths of the two "o" characters in the header.

Feature 8 is the sole that is half-automatically quantified: for each line, the limits of the highest and lowest letters must be entered. Because we have to draw two dots per character and each line has two characters, we request the doctor to mark twenty dots on the monitor to carry out the assessment (Figure 15). The dots are marked, while the required number of dots is shown on the right.

At the conclusion of this process, features 1, 2, 3, 4, 5, 6, 8, 9 and 11 will be automatically computed by the system

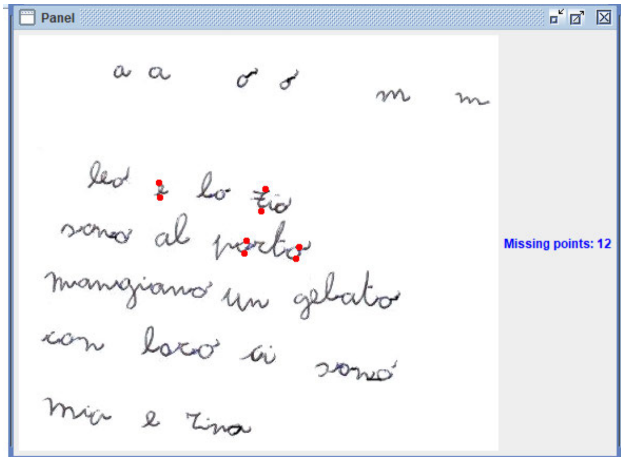

FIGURE 15. Twenty points must be manually set. In the example, eight points have been defined, and twelve points have to be defined.

using the algorithms described in the following sections, and the remaining features must be entered using the dashboard shown in Figure 17. Well-known document analysis algorithms have been adopted (e.g., to segment rows and words); however, some algorithms have been customized for this specific purpose, such as the discovery of skewed writing and the non-aligned left margin.

A. ALGORITHMS

A fundamental activity to start the evaluation of each feature is line and word segmentation. In the literature, the effectiveness of many techniques depends on the type of text.

Segmentation of unbound hand-written lines of text is difficult due to the variability of the line spacing and the variability of the baseline slope. The components of two consecutive lines of text can be touched or superimposed in unbound handwritten text. These superimposed or touching characters greatly complicate the task of segmenting the lines. Due to the specific rules and the text, a simple algorithm that is based on horizontal and vertical histograms on the pixel matrix is sufficient for our scope. The vertical histogram separates words, while the horizontal histogram separates lines. The global horizontal projection method calculates the sum of all black pixels on each line and constructs the corresponding histogram. Based on the peak/valley points of the histogram, the individual lines are segmented. This method has some drawbacks in the case of oblique texts and critical overlapping situations [20].

1) WRITING SIZE ALGORITHM

The average of the letter heights can be determined by considering the first row of text added to the standard text. With the known coordinates of the endpoints of each character, the height will be determined by the variation in the endpoint ordinates. Once the average size is estimated, it is rounded-off to an integer: Table 2 is utilized to obtain the score of feature 1.

2) NON-ALIGNED LEFT MARGIN ALGORITHM

The left margin is estimated by taking into account the first column of words nearest to left border with the exception of the first row. If the spacing in the x-axis from the left margin increases as you scroll through the lines, the left margin is not straight. To derive the value, we determine the difference between the x coordinate of the start of the word in the second row and that of the start of the first word in the last row.

3) SKEWED WRITING ALGORITHM

A computationally efficient procedure for detecting skew lines in scanned documents is based on the cross-referencing of pixels in the vertical lines of a document. Due to the skew, each horizontal text line intersects a predefined set of vertical lines in non-horizontal positions. Using the pixels in these vertical lines, we construct a matrix and evaluate the angle of inclination of the document with considerable accuracy [21]. We adapted this algorithm to our special purposes.

The skewed writing algorithm considers two distinct elements: skewed word and skewed line. If a line is oblique or contains an oblique word, a value of 1 is assigned; otherwise, a value of 0 is assigned. The features score is the sum of the line values. To perform this operation, a vector of minima is created from the dot matrix of the word. The vector of minima is an array that is sized as the length of the word, and for each cell, the height of the first black pixel starting from the bottom is considered. To verify whether the word direction is consistent, the distance between the first minimum of the word m_0 and the mean minimum of the word is determined. A word with a consistent direction has a gap below a threshold. This condition is the formal condition that each word must follow:

$$\forall i \in \{1, \ldots, n\}$$
$$\left| m_0 - \frac{1}{n} \sum_{i=1}^{n} m_i \right| < S_0$$

where n is the number of horizontal pixels in a word, m_i is the minimum in position i and S_0 is the threshold. This threshold and all other thresholds have been empirically defined. However, they can be easily modified by doctors if necessary, especially if the thresholds must be personalized in the cases of frequent monitoring of some patients, based on age, sex or severity of the disease [22]. At the line level, an expanded minima vector that takes into account all the words in the line is formed. From this vector, the mean minimum values of all words can be computed. The gap between two mean values among the words can be calculated to determine if the baseline is coherent. A line with a consistent direction has a distance between the average values below a threshold. In simple terms, the following condition has to be respected for each line:

$$\forall i \in \{1, \ldots, n\} \vee j \in \{1, \ldots, n\} \wedge i \neq j$$
$$\left| \frac{1}{l_i} \sum_{i=1}^{l_i} m_i - \frac{1}{l_j} \sum_{i=1}^{l_j} m_j \right| < S_1$$

where n is the number of horizontal pixels in a line, l_i is the number of pixels of the word in position i and S_1 is the threshold.

At the line level, the first and last minimum should not be too far apart. The absolute value of the distance between the lowest point of the first word and the lower point of the last word is determined by the minimum expanded vector. A line with a consistent pattern has a gap between the two minima under a threshold, which can be edited as previously shown. For each line,

$$|m_0 - m_n| < S_2$$

where n is the number of horizontal pixels in a line, and S_2 is the threshold.

4) INSUFFICIENT SPACE BETWEEN LETTERS ALGORITHM

The dimension of the width of the character ''o'' can be identified by computing the mean of the width (difference between the final abscissa and the initial abscissa) of the two ''o'' characters in the first line. For each line, we can verify if two words exist for which the mean width of the ''o'' is less. The gap between the two words w and $(w+1)$ is the difference between the initial abscissa of the word $(w+1)$ and the final abscissa of the word w.

The feature score is the number of lines in which two words are considerably smaller than the mean width of the ''o''.

5) SHARP ANGLES

As we previously indicated, the sharp angles are sudden peaks upwards or downwards of the text. To identify these angles, we utilize the vector of minima for the previously single words. We verify whether a peak exists by analyzing all minima. If the distance between two minima overcomes a certain threshold, we obtain a sharp angle.

If at least one sharp angle in a line is identified, a value 1 will be assigned to the line; otherwise, a value of 0 is assigned. The final score will be the sum of the line values.

6) BROKEN LINKS BETWEEN LETTERS

To count the number of broken links, we use known information in the text, such as the number of words to segment. If the number of words that are counted exceeds the number of expected words that are counted, the system checks if two groups of characters are so close together that they can be considered the same word but with an interrupted link. In this case, the system checks on which line this link occurs and assigns a score 1 of to the line. The final value will be the number of lines in which at least one broken link is present.

7) IRREGULAR SIZE OF LETTERS ALGORITHMS

As previously mentioned, for this feature, doctor intervention is necessary: he should indicate, for each line, the extremes of the highest and lowest elongation-free characters. The system analyzes the dots pairs per line and computes the differences between two Cartesian ordinates. Recognizing which of the two letters is higher for each line, the system uses the reference sizes reported in table 1 to assign a value of 0 or 1. The feature score is the number of rows where the difference between the highest letter and lowest letter is incoherent.

8) INCONSISTENT HEIGHT BETWEEN LETTERS WITH AND WITHOUT EXTENSION ALGORITHM

To quantify this parameter, we need to determine if each word contains any elongation. A word is considered ''with elongations'' if it contains one of the following letters: y, t, q, p, l, k, j, h, g, f, d, and b.

For each row, the height of the words with elongations and the height of the words without elongations are considered. If at least one ratio between words with elongations and words without elongations overcomes a threshold, a value of 1 will be assigned. The final value is the sum of the row values. For this parameter, the threshold can be chosen by the specialist.

9) TRACED LETTERS

A probabilistic method is used to identify the traced letters. A letter is likely to be traced if it was written darker than the other letters. Once a threshold has been established, the system checks whether a significant stroke of the pen (longer than a certain number of pixels) has been written with greater intensity. By analyzing rows and columns of the pixel matrix, we detect whether too many pixels are darker than the other pixel. A line is assigned a value of 1 if it contains traced strokes; otherwise, a value of 0 is assigned. The final score will be the sum of the line values.

B. TEXT ANALYSIS

After processing the scanned image, a window similar to the window shown in Figure 16 appears to verify the correct segmentation of the text. At the end of the process, a check panel (Figure 17) enables the doctor to validate and approve the results produced by the automatic process.

FIGURE 16. Line and word segmentation check panel.

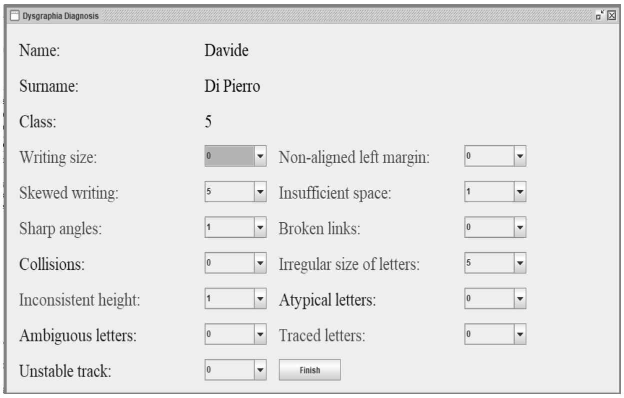

FIGURE 17. Check panel with score results.

As a medical record, a final report is produced and arranged to include all relevant information that a doctor may require, including anamnesis and other personal data.

Reports can be automatically transferred to a patient's personal health record and stored to assess the progression of the disease over time [23]. The first part of the report contains details of the setups and abnormalities for each feature; the second part of the report shows the feature scores and the final mark against the standard limits (upper, lower or normal).

With the involvement of some children who are affected by this disorder, the efficiency of the method and system is presented. As discussed in the next section, the results are interesting. The system is robust in the cases of well separated writing (which is mandatory in BHK) but has minor issues regarding the separation of words in the case of overlapping words.

C. COHORT AND TEST

To validate the effectiveness of the system here presented, we collected handwritten manuscripts from 109 children from the second grade to the fifth grade of primary school who were equally distributed for gender. The writing process was demonstrated for the educators who assisted the children with text writing. Five manuscripts were not acceptable for the test due to missing headers, capital letters or nonconformity with the supplied default text and were rejected.

Table 3 groups our sample according to the school cycle. In Italy, children start attending primary school normally when they are 6. Average age is 8.84 ± 0.94. Our sample includes 12 dysgraphic people over 104.

A health professional manually analyzed the texts following the official protocol. The same process was performed using the software system.

TABLE 3. Children per cycle.

2^{nd}	7
3^{rd}	34
4^{th}	31
5^{th}	32

TABLE 4. Results of expert calculation of features 1, 2, 3, 4, 5, 6, 8, 9 and 11 for each child.

Child grade	P 1	P 2	P 3	P 4	P 5	P 6	P 8	P 9	P 11	Outcome
2	0	0	1	1	4	0	3	1	0	N
2	1	0	0	5	3	0	4	1	0	N
2	0	0	0	5	4	0	3	0	1	N
2	1	0	3	5	4	0	2	0	1	N
2	2	0	0	0	4	0	5	2	0	N
2	2	0	3	4	4	0	5	3	2	N
2	2	0	3	4	5	2	4	1	1	P
3	2	0	3	4	2	3	4	2	0	P
3	1	1	3	4	0	1	2	0	0	N
3	2	0	2	2	4	0	4	0	1	P
3	2	0	2	4	1	0	4	1	0	N
3	3	0	1	2	4	0	5	1	0	N
3	2	0	2	5	5	1	3	1	1	P
4	4	0	4	5	3	2	5	3	3	P
5	2	0	4	5	4	0	4	4	0	P
5	2	0	2	5	2	1	5	1	0	P

TABLE 5. Results of calculation of features 1, 2, 3, 4, 5, 6, 8, 9 and 11 performed by TestGraphia for each child.

Child grade	P 1	P 2	P 3	P 4	P 5	P 6	P 8	P 9	P 11	Outcome
2	0	0	1	1	4	0	3	1	0	N
2	1	0	0	5	5	0	4	0	0	N
2	0	0	1	5	5	0	3	0	2	N
2	0	0	4	5	5	0	2	0	1	N
2	1	0	0	0	4	0	5	2	0	N
2	2	0	3	3	4	0	5	2	1	N
2	2	0	3	3	4	0	5	1	1	P
3	3	0	3	3	2	3	5	2	0	P
3	1	0	2	4	1	0	2	0	0	N
3	2	0	2	2	3	0	3	0	1	P
3	2	0	2	4	3	0	4	1	0	N
3	3	0	1	2	4	0	5	2	0	N
3	3	0	2	4	5	1	3	3	1	P
4	4	0	4	5	4	0	4	4	3	P
5	2	0	4	5	1	0	5	4	0	P
5	3	0	4	5	3	0	3	1	0	P

TABLE 6. Score difference between Table 4 and Table 5.

Child grade	P 1	P 2	P 3	P 4	P 5	P 6	P 8	P 9	P 11
2	0	0	0	0	0	0	0	0	0
2	0	0	0	0	2	0	0	-1	0
2	0	0	1	0	1	0	0	0	1
2	-1	0	1	0	1	0	0	0	0
2	-1	0	0	0	0	0	0	0	0
2	0	0	0	-1	0	0	0	-1	-1
2	0	0	0	-1	-1	-2	1	0	0
3	1	0	0	-1	0	0	1	0	0
3	0	-1	-1	0	1	-1	0	0	0
3	0	0	0	0	-1	0	-1	0	0
3	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
3	1	0	0	-1	0	0	0	1	0
4	0	0	0	0	1	-2	-1	0	0
5	0	0	0	0	-3	0	1	0	0
5	1	0	2	0	1	-1	-2	0	0

IV. RESULTS AND DISCUSSION

The thirteen parameters that were previously detailed were calculated twice: via TestGraphia and manually with assistance from a health professional. In Tables 4 and 5, as an example, the scores calculated for each parameter for

TABLE 7. Differences between the two diagnoses.

TestGraphia		Expert		
BHK Score		BHK Score		
1,13	N	1,13	N	
-0,57	N	-0,88	N	
-0,27	N	-0,11	N	
0,05	N	-0,11	N	
-0,88	N	-1,19	N	
-2,73	P	-2,73	P	
-1,5	P	-1,19	N	FP
-1,5	P	-1,65	P	
-1,34	N	-1,5	P	FN
-0,73	N	-0,73	N	
-0,42	N	-0,57	N	
0,51	N	-0,11	N	
-0,11	N	-0,27	N	
0,97	N	1,74	N	
0,82	N	0,97	N	
0,2	N	0,05	N	
0,51	N	0,36	N	
-1,5	P	-1,5	P	
-0,57	N	-0,42	N	
1,28	N	1,44	N	
0,36	N	0,36	N	
-0,57	N	-0,42	N	
-0,11	N	-0,27	N	
0,36	N	0,51	N	
0,2	N	0,2	N	
1,44	N	1,13	N	
0,05	N	-0,73	N	
0,36	N	0,36	N	
1,13	N	0,97	N	
-0,57	N	0,51	N	
1,28	N	1,28	N	
-0,57	N	-0,42	N	
0,05	N	0,67	N	
0,05	N	0,05	N	
1,28	N	1,59	N	
1,44	N	1,44	N	
-0,27	N	-0,73	N	
0,2	N	0,2	N	
0,51	N	0,36	N	
-0,11	N	-0,11	N	
-1,5	P	-1,19	N	FP
0,51	N	0,67	N	
0,51	N	0,67	N	
-0,57	N	0,05	N	
-0,42	N	-0,11	N	
-0,11	N	-0,27	N	
0,2	N	0,51	N	
-4,11	P	-4,11	P	
0,36	N	0,82	N	
-0,73	N	-0,42	N	
0,97	N	0,36	N	
-1,04	N	-1,19	N	

TestGraphia		Expert		
-0,11	N	-0,42	N	
0,82	N	0,82	N	
0,2	N	-0,27	N	
-2,11	P	-2,11	P	
0,36	N	0,36	N	
0,67	N	0,67	N	
0,2	N	0,2	N	
0,05	N	0,2	N	
-0,27	N	-0,73	N	
-1,19	N	-1,65	P	FN
-0,73	N	-0,73	N	
0,51	N	0,82	N	
0,67	N	0,36	N	
1,44	N	0,97	N	
-0,11	N	0,05	N	
-0,42	N	0,05	N	
0,05	N	0,36	N	
-2,11	P	-1,65	P	
-1,04	N	-0,88	N	
-1,5	P	-1,8	P	
-0,57	N	-0,27	N	
-0,42	N	-0,73	N	
-0,11	N	0,05	N	
-0,57	N	-0,42	N	
0,05	N	0,36	N	
0,51	N	0,67	N	
0,05	N	0,2	N	
0,2	N	0,51	N	
-1,65	P	-1,8	P	
2,05	N	1,9	N	
1,44	N	1,59	N	
0,05	N	0,51	N	
-0,57	N	-0,73	N	
0,97	N	0,51	N	
0,36	N	0,51	N	
1,13	N	0,82	N	
0,51	N	0,82	N	
1,13	N	1,28	N	
0,97	N	1,9	N	
-0,73	N	-0,42	N	
-0,11	N	0,2	N	
-0,27	N	0,36	N	
0,05	N	0,67	N	
-1,19	N	-0,27	N	
-2,57	P	-1,65	P	
0,2	N	1,74	N	
0,67	N	1,74	N	
-1,5	P	-1,65	P	
0,2	N	0,97	N	
-0,11	N	-0,27	N	
-0,73	N	-0,57	N	
0,82	N	0,82	N	

16 children are listed: P1, P2 etc. represent Parameter 1, 2 etc. Outcome is Negative/Positive, yielded taking into consideration all the parameters.

We also compared these scores to show the correspondence between the manually detected features and those determined by the system. As it can be seen in Table 6 most of the difference values are 0; some differences occur but we must take into account that an error in one or two parameters does not cause false results in most cases because the final score is based on 13 parameters and some redundancy is assured, as the specialists affirm.

The most important result to consider is the result obtained by computing the final outcome S by the following formula, which returns the final score for each child, as indicated in the BHK method:

$$S = \frac{M - \sum_{i=1}^{13} P_i}{d} = \frac{M}{d} - \frac{\sum_{i=1}^{13} P_i}{d}$$

where S is the final score, M is 19.3, and d is 6.5, as shown in [19].

The final scores obtained by applying the previous formula are fully reported in Table 7. This table shows the results calculated by the expert and TestGraphia. Table 7 also shows the differences between the two diagnoses: note the presence of 2 false positives, which in any case does not constitute a problem, and 2 false negatives. We remind that most of the software-based diagnostic systems are used as diagnostic support: the doctor is responsible for the validation of the outcome, while software is an extremely important aid, as demonstrated in this case.

System performance must be considered satisfactory, as it can be noted in the confusion matrix in Table 8, while its effectiveness can be summarized as follows:

- *Sensitivity* = *0.83*
- *Specificity* = *0.98*
- *Accuracy* = *0.96.*

For True Negative we mean healthy people predicted as healthy by our system, while for True Positive we mean the dysgraphic people which have been correctly classified as dysgraphic.

Further useful considerations could be done by introducing a further indicator of difference which for this experiment makes the idea of the reliability of the measurement of the software compared to a specialist. Specifically, we can estimate the average difference, in absolute value, between the final score determined by this software and the one calculated by the specialist:

$$M = \frac{1}{N} \sum_{i=1}^{N} |H_i - T_i|$$

TABLE 8. Confusion matrix.

Predicted/True Condition	Positive	Negative
Positive	10	2
Negative	2	90

where M is the average difference, H_i is the i-th score value determined by the specialist and T_i the i-th score determined by TestGraphia. In our case, $M = 0.28 \pm .28$ means that, in the worst case, children that have a score of -1.77 ($-1.77+0.28 = -1.49 \rightarrow$ Negative) are not recognized as dysgraphic; indeed -1.77 actually means that the level of dysgraphia is not high: most serious cases reveal a score <-2. These last cases corresponding to border values could be labeled as 'suspect cases'.

Some drawbacks are admitted in our system due to the fact that the borderline subjectivity is not fully removable also in the case of expert diagnosis, as BHK itself shows a certain ambiguity. There may be also some errors due to millimeter variations. On the other hand, it is also possible that a significant variation is considered as an error by the software while for an expert it is an alternative way for a child to denote a letter. The slight variability of the parameters is also present among the experts themselves, due to the difficulty in providing an objective evaluation on writing, specially on children's writing. In this context, the main focus of TestGraphia is to reduce as much as possible ambiguity, aiming at the more reliable interpretation of anomalies.

V. CONCLUSION

This research and software have enabled the design of a support system for doctors who arrange systematic screenings and widespread execution of BHK tests: these tests can relieve their tiring and boring activities and conserve resources while reducing costs. The main goal is to detect early symptoms and examine a larger number of children in a shorter time.

Well-known document analysis algorithms have been employed in this research. However, some of these algorithms have been modified and adapted for this specific purpose, such as the skewed writing algorithm and the non-aligned left margin algorithm.

We suppose that TestGraphia, thanks to its ease of use and no cost, could be an interesting opportunity to facilitate the diagnosis and raise awareness of these problems.

It allows a reliable diagnosis even with an asynchronous relationship between patient and doctor, who can examine the results without attending the test. It takes into account the expertise of the doctor which is important to validate the final result, while other devices-based systems are seen as black boxes by specialists. Also privacy is preserved in carrying out the test, which can also be done at home without using special devices.

Finally, it is important to mention the fact that the techniques and devices presented in other publications are not used at all in Italy and probably they are not widespread in the world.

At this time, we consider the system performance to be satisfactory. It will certainly be interesting to continue to study whether and what improvements can be achieved using different technologies, including machine learning.

We are studying the efficacy of applying digital tools based on the dynamic analysis of writing, introducing other parameters such as pressure, and not relying only on the handwritten text; we are also interested in verifying if the use of different writing tools (as special devices) can introduce a bias in the final result: it is evident that, for example, the friction of the pen on the sheet is a significant factor that can affect the amplitude of the movements of a child's hand.

ACKNOWLEDGMENT

The authors are grateful to the health professional Dr. M. Bovino for her assistance and collaboration. Furthermore, the authors are grateful to Edizioni Centro Studi Erickson S.p.A. (Trento, Italy) for granting permission to partially reproduce in this article, the material included in [27].

REFERENCES

[1] M. Pratelli, *Dysgraphia*. Portland, OR, USA: Erickson, 1995.
[2] S. Rosenblum, P. L. Weiss, and S. Parush, "Handwriting evaluation for developmental dysgraphia: Process versus product," *Reading Writing*, vol. 17, p. 433, Jul. 2004, doi: 10.1023/B:READ.0000044596.91833.55.
[3] R. K. Deuel, "Developmental dysgraphia and motor skills disorders," *J. Child Neurol.*, vol. 10, no. 1, pp. S6–S8, Jan. 1995, doi: 10.1177/08830738950100s103.
[4] E. D. Dmitrova, N. V. Dubrovinskaya, I. P. Lukashevich, "Features of cerebral support of verbal processes in children with dysgraphia and dyslexia," *Hum. Physiol.*, vol. 31, p. 125, Mar. 2005, doi: 10.1007/s10747-005-0020-5.
[5] T. T. S. Ingram, "Paediatric aspects of specific developmental dysphasia, dyslexia and dysgraphia," *Develop. Med. Child Neurol.*, vol. 2, no. 4, pp. 254–277, Jul. 2010.
[6] E. Adijapha, Y. Landau, L. Frenkel, M. Teicher, V. Grosstsur, and R. Shalev, "ADHD and dysgraphia: Underlying mechanisms," *Cortex*, vol. 43, no. 6, pp. 700–709, 2007, doi: 10.1016/s0010-9452(08)70499-4.
[7] J. Mortley, P. Enderby, and B. Petheram, "Using a computer to improve functional writing in a patient with severe dysgraphia," *Aphasiology*, vol. 15, no. 5, pp. 443–461, May 2001, doi: 10.1080/02687040042000188.
[8] J. Danna, "Handwriting movement sonification for the rehabilitation of dysgraphia," in *Proc. 10th Int. Symp. Comput. Music Multidisciplinary Res.*, Marseille, France, Oct. 2013, pp. 200–208.
[9] G. Dimauro, V. Di Nicola, V. Bevilacqua, D. Caivano, and F. Girardi, "Assessment of speech intelligibility in parkinson–disease using a speech-to-text system," *IEEE Access*, vol. 5, pp. 22199–22208, 2017, doi: 10.1109/access.2017.2762475.
[10] C. Loconsole, G. D. Cascarano, A. Brunetti, G. F. Trotta, G. Losavio, V. Bevilacqua, and E. Di Sciascio, "A model-free technique based on computer vision and sEMG for classification in Parkinson–disease by using computer-assisted handwriting analysis," *Pattern Recognit. Lett.*, vol. 121, pp. 28–36, Apr. 2019, doi: 10.1016/j.patrec.2018.04.006.
[11] G. Dimauro, D. Caivano, V. Bevilacqua, F. Girardi, and V. Napoletano, "Voxtester, software for digital evaluation of speech changes in parkinson disease," in *Proc. IEEE Int. Symp. Med. Measurements Appl. (MeMeA)*, Benevento, Italy, May 2016, pp. 1–6, doi: 10.1109/memea.2016.7533761.
[12] M. C. Silveri, S. Misciagna, M. G. Leggio, and M. Molinari, "Spatial dysgraphia and cerebellar lesion," *Neurology*, vol. 48, no. 6, pp. 1529–1532, Jun. 1997, doi: 10.1212/WNL.48.6.1529.
[13] J. Danna, "Handwriting sonification for the diagnosis of dysgraphia," in *Proc. 10th Int. Symp. Comput. Music Multidiscip. Res. Sound, Music Motion*, Feb. 2013, pp. 123–126.
[14] A. Caute and C. Woolf, "Using voice recognition software to improve communicative writing and social participation in an individual with severe acquired dysgraphia: An experimental single case therapy study," *Aphasiology*, vol. 30, nos. 2–3, pp. 245–268, 2016, doi: 10.1080/02687038.2015.1041095.
[15] C. Estes and R. L. Bloom, "Using voice recognition software to treat dysgraphia in a patient with conduction aphasia," *Aphasiology*, vol. 25, no. 3, pp. 366–385, Feb. 2011.
[16] B. E. Al, "Evaluation of handwriting movement kinematics: From an ecological to a magnetic resonance environment," *Front. Hum. Neurosci.*, vol. 10, p. 488, Sep. 2016, doi: 10.3389/fnhum.2016.00488.
[17] V. Bevilacqua, A. Brunetti, G. F. Trotta, G. Dimauro, K. Elez, V. Alberotanza, and A. Scardapane, "A novel approach for Hepatocellular Carcinoma detection and classification based on triphasic CT Protocol," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, San Sebastian, Spain, Jun. 2017, pp. 1856–1863, doi: 10.1109/cec.2017.7969527.
[18] G. Dimauro, "A new image quality metric based on human visual system," in *Proc. IEEE Int. Conf. Virtual Environ. Hum.-Comput. Inter. Meas. Syst. (VECIMS)*, Tianjin, China, Jul. 2012, pp. 69–73 doi: 10.1109/vecims.2012.6273216.
[19] E. Hamstra-Bletz, B. J. De, and B. P. L. M. Den Brinker, *BHK—Concise Evaluation Scale for Children's Handwriting*. Lisse, The Netherlands: Swets & Zeitlinger, 1987.
[20] U. Pal and S. Datta, "Segmentation of Bangla unconstrained handwritten text," in *Proc. 7th Int. Conf. Document Anal. Recognit.*, Edinburgh, U.K., Apr. 2005, pp. 1128–1132, doi: 10.1109/icdar.2003.1227832.
[21] B. Gatos, N. Papamarkos, and C. Chamzas, "Skew detection and text line position determination in digitized documents," *Pattern Recognit.*, vol. 30, no. 9, pp. 1505–1519, Sep. 1997.
[22] D. Caivano, "Personal health E-record–toward an enabling ambient assisted living technology for communication and information sharing between patients and care providers," in *Ambient Assisted Living* (Lecture Notes in Electrical Engineering), vol. 544. Cham, Switzerland: Springer, 2018, doi: 10.1007/978-3-030-05921-7.
[23] G. Dimauro, D. Caivano, F. Girardi, and M. M. Ciccone, "The patient centered Electronic Multimedia Health Fascicle–EMHF," in *Proc. IEEE Workshop Biometric Meas. Syst. Secur. Med. Appl. (BIOMS)*, Rome, Italy, Oct. 2014, pp. 61–66, doi: 10.1109/bioms.2014.6951537.
[24] J. Mekyska, M. Faundez-Zanuy, Z. Mzourek, Z. Galaz, Z. Smekal, and S. Rosenblum, "Identification and rating of developmental dysgraphia by handwriting analysis," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 2, pp. 235–248, Apr. 2017.
[25] T. Asselborn, "Automated human-level diagnosis of dysgraphia using a consumer tablet," *NPJ Digit. Med.*, vol. 1, no. 1, p. 42, Aug. 2018.
[26] S. Rosenblum and G. Dror, "Identifying developmental dysgraphia characteristics utilizing handwriting classification methods," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 2, pp. 293–298, Apr. 2017.
[27] C. Di Brina and G. Rossini, *Test BHK—Scala Sintetica per la Valutazione Della Scrittura in Età Evolutiva*. Portland, OR, USA: Erickson, 2011.

GIOVANNI DIMAURO was born in Taranto, Italy, in 1964. He received the Laurea degree in computer science from the Università degli Studi di Bari Aldo Moro, in 1987, where he is currently an Associate Professor with the Computer Science Department. He teaches computer programming and multimedia systems. His research interests include e-health, multimedia systems, and pattern recognition with applications in medicine, such as new diagnosis technology for anemia and Parkinson's disease. He is the author of more than 150 articles, which have been published in scientific journals, proceedings, and books. He is the author of two patents, of which the latter pertains to the field of noninvasive anemia estimation. For additional details visit: http://www.di.uniba.it/∼dimauro/.

VITOANTONIO BEVILACQUA is currently an Associate Professor of electronic and information bioengineering with the Department of Electrical and Information Engineering, Polytechnic University of Bari, Italy. Since 1996, he has been working and investigating in the field of computer vision and medical image processing, human–machine interaction, bioengineering, machine learning, and soft computing. In May 2019, he joined the National Scientific Habilitation (ASN), as a Full Professor of information processing systems, and has been the Head of the Industrial Informatics Lab, since April 2017. Since March 2019, he has been the CEO of Apulian Bioengineering, which is a spin-off and innovative start-up company at the Department of Electrical and Information Engineering, Polytechnic University of Bari. He has published more than 180 articles in refereed journals, books, and international conferences proceedings.

LUCIO COLIZZI received the degree from the University of Salento. After graduating from the University of Salento, he worked in the private sector for the company Infobyte Spa in the field of design and development of virtual reality applications. In 2001, he became the Director of the Computer Engineering Division of Technologies Design and Materials European Research Centre (CETMA) and directs more than 30 large industrial research projects. He is currently an Information Technology Engineer. He is the CEO of the innovative startup Smart Education Lab, which is a spinoff of CETMA that operates in the field of digital fabrication, the IoT, and process digitalization and simulation. He is the author of a European patent in the field of mechatronics and several scientific publications. His areas of expertise are focused on the development of ICT technologies in Industry 4.0, cultural heritage, and biomedical robotics.

DAVIDE DI PIERRO was born in Bari, Italy, in 1997. He received the Laurea degree in computer science and software production technologies from the Università degli Studi di Bari Aldo Moro, in 2019, where he is currently pursuing the master's degree in computer science.