

A Machine Learning Algorithm Applied to Trichoscopy for Androgenic Alopecia Staging and Severity Assessment

Marco Di Fraia¹, Lorenzo Tieghi², Francesca Magri¹, Gemma Caro¹, Simone Michelini¹, Giovanni Pellacani¹, Alfredo Rossi¹

¹ Dermatology Clinic, Department of Clinical Internal, Anesthesiologic and Cardiovascular Sciences, Sapienza University of Rome, Roma, Italy

² Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Roma, Italy

Key words: androgenic alopecia, trichoscopy, machine learning, support vector machine, staging

Citation: Di Fraia M, Tieghi L, Magri F, et al. A Machine Learning Algorithm Applied to Trichoscopy for Androgenic Alopecia Staging and Severity Assessment. *Dermatol Pract Concept*. 2023;13(3):e2023136. DOI: <https://doi.org/10.5826/dpc.1303a136>

Accepted: December 7, 2022; **Published:** July 2023

Copyright: ©2023 Di Fraia et al. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (BY-NC-4.0), <https://creativecommons.org/licenses/by-nc/4.0/>, which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original authors and source are credited.

Funding: None.

Competing Interests: None.

Authorship: All authors have contributed significantly to this publication.

Corresponding Author: Di Fraia Marco, Dermatology Clinic, Department of Clinical Internal, Anesthesiologic and Cardiovascular Sciences, Sapienza University of Rome, Viale del Policlinico 155, 00161, Roma (RM), Italy. Email: marco.difraia@uniroma1.it

ABSTRACT **Introduction:** Androgenic alopecia (AGA) staging is still based on macroscopic scales, yet the introduction of trichoscopy is gradually bringing an important change, even though it remains an eye-based method. However, recently developed artificial intelligence-assisted programs can execute automated count of trichoscopic patterns. Nevertheless, to interpret data elaborated by these programs can be complex. Machine learning algorithms might represent an innovative solution. Among them, support vector machine (SVM) models are among the best methods for classification.

Objectives: Our aim was to develop a SVM algorithm, based on three trichoscopic patterns, able to classify AGA patients and to calculate a severity index.

Methods: We retrospectively analyzed trichoscopic images from 200 AGA patients using Trichoscale Pro® software, calculating the number of vellus hair, empty follicles and single hair follicular units. Then, we elaborated a SVM model, based on these three patterns and on sex, able to classify patients as affected by mild AGA or moderate-severe AGA, and able to calculate the probability of the classification being correct, expressed as percentage (from 50% to 100%). This probability estimate is higher in patients with more AGA trichoscopic patterns and, thus, it might serve as a severity index.

Results: For training and test datasets, accuracy was 94.3% and 90.0% respectively, while the Area Under the Curve was 0.99 and 0.95 respectively.

Conclusions: We believe our SVM model could be of great support for dermatologists in the management of AGA, especially in better assessing disease severity and, thus, in prescribing a more appropriate therapy.

Introduction

Androgenic alopecia (AGA) is the most common form of hair loss, affecting around 80% of men and 50% of women in the course of their life [1]. The fundamental role of androgen hormones and genetic susceptibility in AGA pathogenesis has been widely demonstrated [2]. Nowadays, AGA staging is still based on macroscopic scales, like Norwood-Hamilton for men and Sinclair for women [3,4], but the introduction of trichoscopy in scalp examination is gradually bringing an important change. Analysis of trichoscopic patterns, like vellus hairs or single hair follicular units, helps dermatologists better assess disease activity, despite of AGA extension [5]. However, trichoscopy remains an eye-based method. In fact, it is almost impossible to count the exact number of patterns that are visualized per each field of view. This task can be performed by recently developed artificial intelligence-assisted programs for hair analysis, which can execute automated count of hair density, hair diameter, etc [6,7]. Nevertheless, to interpret data generated by these programs can be complex and time consuming. Machine learning algorithms might represent an innovative solution. These models are generated from sample data and their purpose is to make predictions or decisions without being explicitly programmed to do so. Among them, support vector machine (SVM) models are among the best methods for classification [8]. Simply explained, given a set of subjects labelled as belonging to one of two categories (training dataset), SVM algorithms are able to assign new subjects (test dataset) to one of the previous categories.

Objectives

Our objective was to develop a SVM algorithm, based on three trichoscopic patterns, which could classify patients as affected by mild AGA or moderate-severe AGA and could elaborate a severity index for each patient.

Methods

In this single-center observational retrospective study we analyzed trichoscopic images from 200 patients (100 males and 100 females) affected by AGA. Clinical (macroscopic) disease severity was assessed using Norwood-Hamilton scale

for men and Sinclair scale for women. Fifty male patients were affected by mild AGA (20 classified as stage I and 30 as stage II) and 50 by moderate-severe AGA (30 classified as stage IV and 20 as stage V). As for men, women patients were equally affected by mild AGA (50 classified as stage II) and moderate-severe AGA (50 classified as stage IV). At the time of analysis, patients were naive to any treatment for AGA and were not affected by any other scalp disease. Patients images were collected in our outpatient clinic from September 2019 to December 2021.

The scalp of all subjects was combed along the midline to the sides and evaluated using the video-dermoscope Medicam 1000® (FotoFinder Systems GmbH). Digital trichoscopic photographs were obtained in standardized scalp areas according to the patient sex. In women, this area corresponded to the intersection point between the line connecting the ears and the line connecting the tip of the nose to the vertex. In men, the area was localized two centimeters ahead of the point previously described for women. Field of view area was equivalent to 0.903 cm². Photographs were collected at 20-fold magnification.

Each trichoscopic image was analyzed with Trichoscale Pro® software (FotoFinder Systems GmbH), which allows to perform accurate automatic measurements of scalp structures, with subsequent manual verification. In our study, we considered the following trichoscopic patterns:

1. vellus hairs (vellus hairs / total hairs ratio, expressed as percentage);
2. single hair follicular units (single hair follicular units / total follicular units ratio, expressed as percentage);
3. empty follicles (number of empty follicles).

We elaborated a SVM model based on these three patterns and on sex. This model is able to classify each patient as affected by mild AGA or moderate-severe AGA. For each subject, the algorithm also calculates the probability of the classification being correct (from this point on referred to as “probability estimate”), expressed as percentage, ranging from 50% (unreliable classification) to 100% (exact classification).

As common practice for these kinds of models [8], population was split into 70% (N = 140) as training and 30% (N = 60) as test set. For each dataset, we calculated the following performance evaluation metrics: sensitivity (corresponding

to recall), specificity, positive predictive value (corresponding to precision), negative predictive value, accuracy, receiver operating characteristic curve (ROC) curve analysis.

SVM model code is available upon direct request to the authors.

Statistical Analysis

Mann-Whitney U test was used for comparison between trichoscopic parameters. The SVM model is based on open-source codes. The code was developed in a Python 3.8 based framework. The probability estimate is mathematically derived by normalizing the algorithm function by the weights of the hyperplane. Artificial intelligence was trained using the sklearn library.

Results

Medians and interquartile ranges of each trichoscopic parameter, for each study group, are listed in Table 1. All of the three parameters were significantly higher, both in men and women, in patients affected by moderate-severe AGA compared to mild AGA patients (Table 1).

Trichoscopic parameter values for each patient, and the hyperplane function dividing mild AGA patients from moderate-severe ones, are reported in the 3D scatter plot (Figure 1).

The estimates of correct classification, as a function of the distance of the example from the decision boundary, for each patient of the training dataset, are reported in Figure 2.

Table 1. Comparison between trichoscopic parameter values of mild AGA patients and moderate-severe ones (median and interquartile range).

	Mild	Moderate-severe	P
Male (N=100)			
Vellus hairs (%)	27 (21-32)	46.5 (38-63)	< 0.0001
Single hair FU (%)	37 (32-42)	46 (39-35)	< 0.0001
Empty follicles (N)	0 (0-0)	1 (0-2)	< 0.0001
Female (N=100)			
Vellus hairs (%)	33 (27-40)	62 (58-69)	< 0.0001
Single hair FU (%)	45.5 (39-51)	54 (49-59)	< 0.0001
Empty follicles (N)	0 (0-0)	1 (0-3)	= 0.0002

FU = follicular units.

* Mann-Whitney U-Test.

Model performance metrics values, for training and test datasets, are reported in Table 2.

ROC curve analysis, for training and test datasets, are reported in Figure 3.

Conclusions

In recent times, trichoscopy has become an essential tool for the diagnosis of hair loss disorders [5]. The analysis of specific patterns allows dermatologists to perform accurate evaluation of disease severity and activity. However, this process is inevitably subjective and eye based. The developing of new artificial intelligence programs [6,7], capable of precisely quantifying these patterns, opens new scenarios in the field of trichoscopy. In our study, we analyzed trichoscopic images with Trichoscale Pro® software, and performed an accurate manual verification of the calculated values.

In the study population, the percentage of vellus hairs, the percentage of single hair follicular units and the number of empty follicles were significantly higher in patients affected by clinical (using classical macroscopic scales) moderate-severe AGA compared to clinical mild AGA patients (Table 1), as expected according to literature [8]. Nevertheless, 8% of patients diagnosed with clinical mild AGA were classified as affected by moderate-severe AGA patients by our model, which is based on trichoscopic patterns. We presume this might be related to the fact that clinical macroscopic observation alone is not sufficient to determine the real severity and activity of AGA. According to our model, in fact, an increase in trichoscopic parameter values is correlated to a higher probability (expressed as percentage) to be affected by moderate-severe AGA. This probability estimate might serve as a severity index, given that, for patients classified as affected by moderate-severe AGA, the more this value approaches 100%, the higher is the number of AGA trichoscopic findings that are observed in the patient. On the contrary, for patients classified as affected by mild AGA, the more the estimate value approaches 100%, the lower is the number of AGA trichoscopic findings.

Our SVM model might be very useful in alerting dermatologist about those patients presenting with not extensive alopecia, diagnosed as mild AGA according to classical macroscopic scales, but actually affected by moderate-severe AGA, after trichoscopic analysis. As an example, a patient diagnosed with stage II AGA (according to Norwood-Hamilton scale), after Trichoscale Pro® analysis, exhibits 42% vellus hairs, 45% single hair follicular units and 1 empty follicle. Our algorithm classifies this patient as affected by moderate-severe AGA, with a probability estimate of 77% (Figure 4).

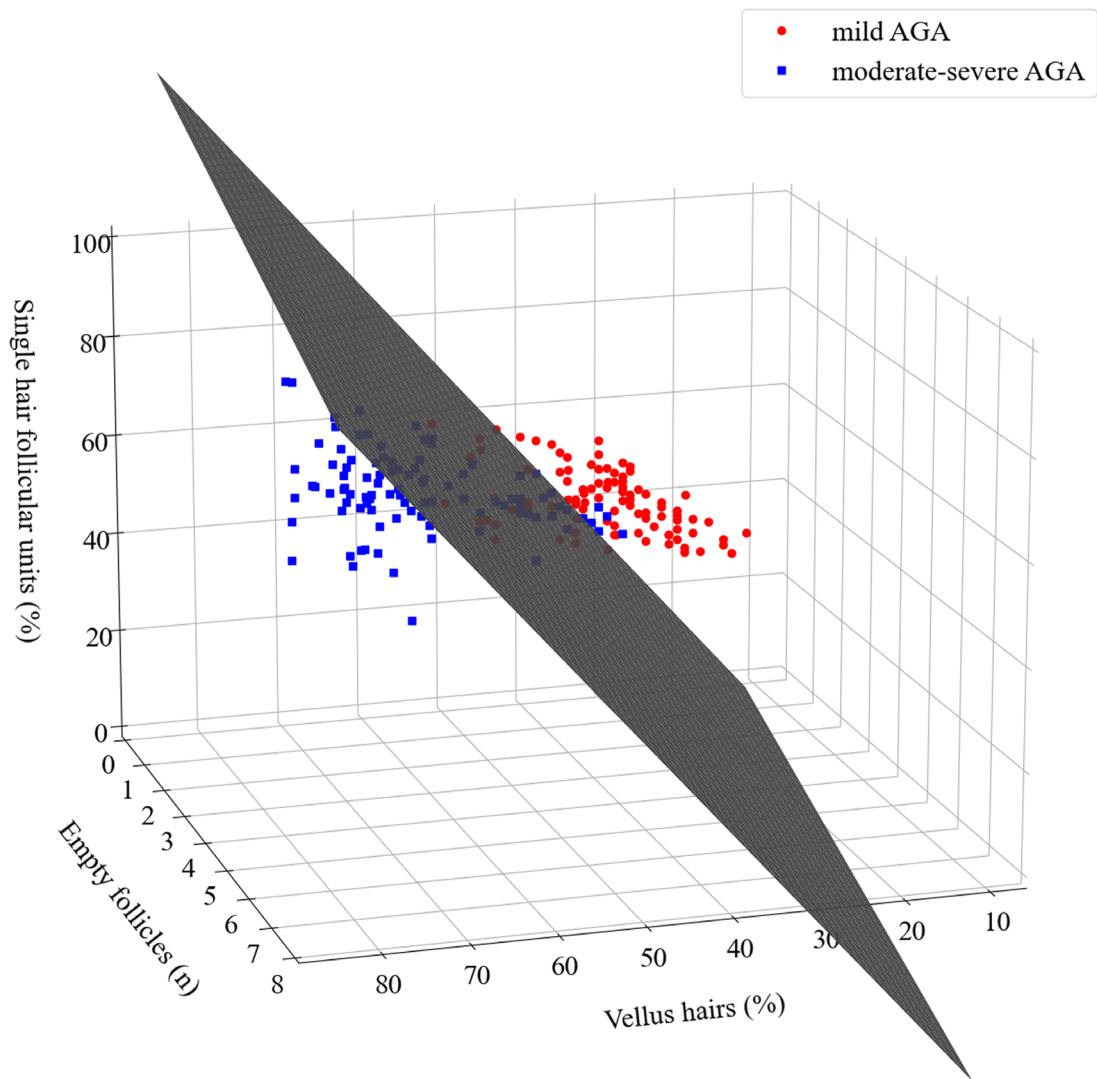


Figure 1. 3D scatter plot showing the three trichoscopic parameter values for each patient. The hyperplane function dividing mild androgenic alopecia (AGA) patients (red) from moderate-severe ones (blue) is colored in gray.

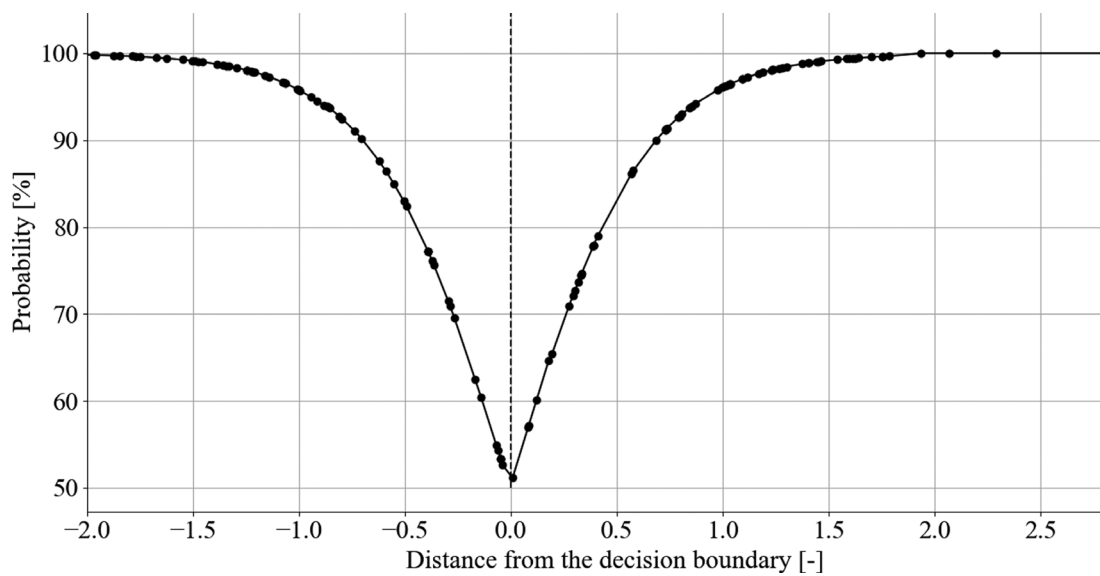


Figure 2. The probability estimates of correct classification, as a function of the distance of the example from the decision boundary, for each patient of the training dataset.

The severity index elaborated by our machine learning model could also guide dermatologists in choosing the right treatment regimen for each patient. Indeed, subjects with a high probability to be affected by moderate-severe AGA might benefit of a more aggressive therapeutic approach. Moreover, it could be useful for monitoring response to therapy: a decreased probability estimate, calculated after a treatment period, would mean that AGA severity is decreasing as well.

The high performance metrics values, like precision, accuracy and recall, demonstrate that the algorithm, in our study population, works extremely well as AGA classifier in the majority of cases. Obviously, the test dataset performance metrics values are lower than the training ones, as expected for this kind of machine learning models in order to avoid overfitting (it occurs when the model fits too closely to the training set, and it is unable to generalize well to new data) [8]. The probability estimates curves slopes (Figure 2)

Table 2. Model performance metrics.

	Train	Test
Sensitivity	94.2%	93.6%
Specificity	94.4%	86.2%
PPV	94.2%	87.8%
NPV	94.4%	92.6%
Accuracy	94.3%	90.0%
AUC	0.99	0.95

AUC = area under the curve; NPV = negative predictive value; PPV = positive predictive value.

are considerably high, and for most patients (81.5%), values are higher than 80%.

Our study has of course some limitations. First of all, Trichoscale Pro® analysis needed manual verification for some minor corrections in most of the cases (78%). This might be attributed to the fact that, in the analyzed scalp areas, hairs were not cut to a short length, even if were accurately combed to reach the best analysis result.

The major algorithm limitation is that, for patients with a probability estimate near 50% (who actually are a small minority of our study population), classification might be unreliable and would need dermatologist careful interpretation. In this regard, we chose not to include stage III AGA patients (according to Norwood-Hamilton or Sinclair scales) because those subjects would have placed close to the algorithm decision boundary and probably would have negatively influenced the model performance. Nevertheless, the algorithm is suitable for all AGA stages, even though, in our opinion, it is best applicable in mild and moderate stages. In fact, the principal aim of our model is to reveal those patients presenting with a macroscopically appearing mild alopecia although having a trichoscopically active disease that might quickly advance to more severe stages.

Naturally, it is necessary to test model's ability to generalize on a wider population sample, enrolling patients from centers familiar with the use of trichoscopic pattern analysis software like Trichoscale Pro®.

Recently, artificial intelligence has made great progresses in the study of hair disorders, and many programs have recently been developed for the automated count of hair number, hair diameter and other trichoscopic patterns [6,7].

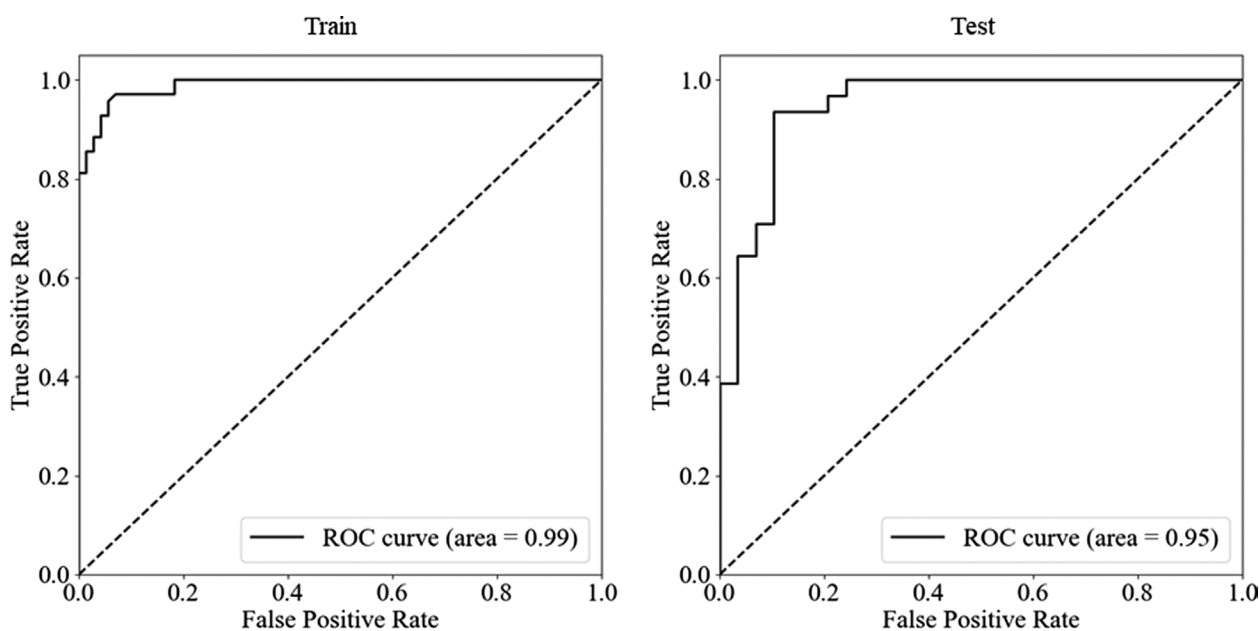


Figure 3. Receiver operating characteristic curve (ROC) curve analysis, for training (left) and test (right) datasets.

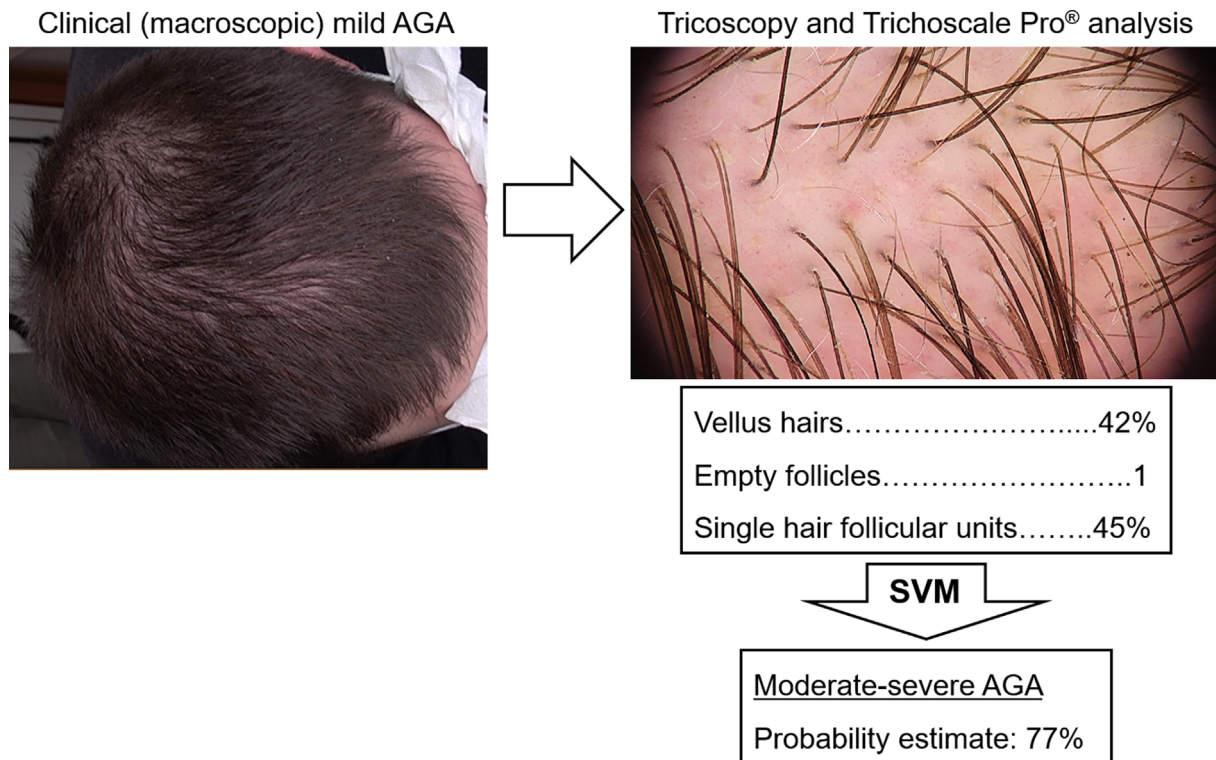


Figure 4. Example of model application. After macroscopic examination, a 20x trichoscopic photograph is taken and trichoscopic patterns are automatically analyzed by Trichoscale Pro® program. The calculated values are then inserted into the support vector machine (SVM) algorithm, which classify the patient as affected by moderate-severe androgenic alopecia (AGA), with a 77% probability of the decision being correct.

However, based upon our knowledge, no algorithm has yet been elaborated to classify patients affected by AGA, neither to assess disease severity. For this reason, we believe our model could be of great support for dermatologists in the management of AGA, especially in defining a more precise staging and, thus, in prescribing a more appropriate therapy.

References

- Alessandrini A, Starace M, D'Ovidio R, et al. Androgenetic alopecia in women and men: Italian guidelines adapted from European Dermatology Forum/European Academy of Dermatology and Venereology guidelines. *G Ital Dermatol Venereol*. 2020;155(5):622-631. DOI: 10.23736/S0392-0488.19.06399-5. PMID: 33295740.
- Lolli F, Pallotti F, Rossi A, et al. Androgenetic alopecia: a review. *Endocrine*. 2017;57(1):9-17. DOI: 10.1007/s12020-017-1280-y. PMID: 28349362.
- Norwood OT. Male pattern baldness: classification and incidence. *South Med J*. 1975;68(11):1359-1365. DOI: 10.1097/00007611-197511000-00009. PMID: 1188424.
- Biondo S, Goble D, Sinclair R. Women who present with female pattern hair loss tend to underestimate the severity of their hair loss. *Br J Dermatol*. 2004;150(4):750-752. DOI: 10.1046/j.0007-0963.2003.05809.x. PMID: 15099373.
- Starace M, Orlando G, Alessandrini A, Piraccini BM. Female Androgenetic Alopecia: An Update on Diagnosis and Management. *Am J Clin Dermatol*. 2020;21(1):69-84. DOI: 10.1007/s40257-019-00479-x. PMID: 31677111.
- Gupta AK, Ivanova IA, Renaud HJ. How good is artificial intelligence (AI) at solving hairy problems? A review of AI applications in hair restoration and hair disorders. *Dermatol Ther*. 2021;34(2):e14811. DOI: 10.1111/dth.14811. PMID: 33496058.
- Jartarkar S, Patil A, Waskiel-Burnat A, Rudnicka L, Starace M, Grabbe S, Goldust M. Artificial Intelligence in Hair and Nail Disorders. *J Drugs Dermatol*. 2022;21(10):1049-1052 DOI: 10.36849/JDD.6519. PMID: 36219051.
- Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273-297. DOI: 10.1007/BF00994018
- Hu R, Xu F, Han Y, Sheng Y, Qi S, Miao Y, Yang Q. Trichoscopic findings of androgenetic alopecia and their association with disease severity. *J Dermatol*. 2015;42(6):602-607. DOI: 10.1111/1346-8138.12857. PMID: 25810236.