INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Centrum teorie vzdělávání přírodovědných oborů
Reg. č.: CZ.1.07/2.3.00/20.0166

Zahraniční konference v rámci projektu ESF

Prof. RNDr. Jaromír Vaňhara, CSc., Katedra biologie PdF UPOL

Cílová instituce*:* Université Libre de Bruxelles, Brusel, Belgie

Termín: 18.-21. říjen 2012

Účelem cesty byla účast na mezinárodní konferenci Benelux Congress of Zoology 2012, která se konala v Bruselu (Université Libre de Bruxelles).

Hlavním cílem pracovní cesty byla prezentace výsledků zoologické části evropského projektu CETPO na Benelux Congress of Zoology 2012. Aktivní prezentace se týkala testování využití umělých neuronových sítí v zoologické systematice. Téma prezentovaného posteru: Vanhara J., M. Bocakova, P. Fedor, K. Janisova, E.M. Peña Mendéz, N. Dvorakova-Murarikova & J. Havel 2012: Biological objects recognition Artificial Neural Networks Identification: mayor step towards integrative taxonomy. Abstract book of 19th Benelux Congress of Zoology, p. 64

19th Benelux Congress of Zoology, Integrative biology: From ecology to molecules, 19-20 October 2012

Biological objects recognition

ARTIFICIAL NEURAL NETWORKS IDENTIFICATION:
a major step towards integrative taxonomy

Vaňhara J. [1,2)], Bocáková M. [2)], Fedor P. [3)], Dvořáková-Muráriková N. [1)], Janišová K. [2)], Peña Mendéz E.M. [4)] & Havel J. [1)]

[1)] Masaryk University, Brno, Czech Republic; vanhara@sci.muni.cz
[2)] Palacký University, Olomouc, Czech Republic
[3)] Comenius University, Bratislava, Slovakia
[4)] University of La Laguna, Tenerife, Spain

Introduction

Artificial Neural Networks (ANN) are statistical tools that can be used as a part of an integrative taxonomy approach, together with morphological, chemical or molecular evaluation. The advantages of ANN include the ability to learn from examples under study and to generalize observed characters. ANN are able to bring solutions not only with the species identification but also with the resolution of some taxonomic problems. Precise and accurate training set is a basis for obtaining reliable identification. The supervised feed-forward ANN with back-propagation of errors was applied on morphometric and qualitative characters, and chemometric data in databases with correctly identified species.

ANN were applied by us on different taxa such as:

- *Patella* limpets (Molusca), Chemosphere 59 (2006), 1038-1048
- thrips (Thysanoptera), Biol. Bull. Fac. St [illegible] 35 (2010), 475-487
- insect pests (Thysanoptera), Plant. Protect. Sci. 46 (2010), 160-168; J. Insect Sci., submitted (2012)
- *Tachina* flies (Diptera), Biologia, Bratislava 65 (2010), 442-449; Biol. Bull. Fac. Sci. 35 (2011), 149-159
- *Libentis* beetles (Coleoptera) and Acrolitha flies (Diptera) are studied by the authors now

Tachina (Diptera)

ANN methodology was applied for the identification of West Palaearctic parasitic flies of the genus *Tachina* (Tachinidae, Diptera). The taxonomic complicacy of the genus can be illustrated through 43 synonyms for 12 currently valid species. ANN also helped to clear several taxonomic problems:

1) the subgeneric classification of *Tachina* was changed (24,5,2),
2) a new Ibero-Alpine species was revealed in the examined material (21,6,3)
3) one synonym was revoked as pertaining to a valid species (28,4,3)
4) one species in central Europe proved to be wrongly identified and was excluded from national checklist (28,4,5)

The correctness were tested by partially performed molecular analyses based on two mitochondrial markers (12S, 16S rDNA) and a detailed comparative morphological examination (male postabdomen structures). This three approaches lead to congruent results (Mareščková et al. 2011, Vanhara et al. 2007).

Tachina sp. (photo F. Fehl)

Morphometric characters used

Tachina ANN optimal architecture (24,5,5), subgenera

Patella (Molusca)

First use of ANN for chemometric identification in marine intertidal and subtidal limpets from Canary Isles (Isl. Fuerteventura). On the base of content of aliphatic hydrocarbons in the soft tissues by gas chromatography, data from *Patella piperata*, *P. candei crenata* and *P. ulyssiponensis aspera* were treated with ANN and it was found that using suitable architecture of a supervised ANN, the limpets can be successfully identified up to 98% (Hernandez-Borges et al., 2004).

Patella ANN architecture (10,6,1)

Patella ulyssiponensis aspera

Libentis (Coleoptera)

Comparative morphology approach is usually used for the species identification of beetles. For ANN methodology, 17 morphometric characters of 66 identified specimens of *Libentis* (Lycidae) formed a database, which was applied to identification of *Libentis* males for the first time. When ANN multivariate approach was used to distinguish unidentified specimens, several species new to science were confirmed (unpublished study).

Thrips (Thysanoptera)

Training database includes morphological characters recorded for 208 specimens of common European Thysanoptera belonging to 18 species and four genera (*Aeolothrips*, *Chirothrips*, *Dendrothrips*, and *Limothrips*). Using 20 input characters (different lengths measured on thrips body, presence/absence characters and sex) and a single supervised ANN (multilayer perceptron) architecture we reached 97 % correct identification of species. Optimal architecture (20,7,18) . (Fedor et al. 2008).

Possible thrips ANN architecture (10,6,5,18)

Morphometric characters used for thrips

Pest thrips (1-2 mm)

Libentis ANN architecture (17,10,26)

Libentis sp. (orig. M. Bocáková)

Agriculture pests

Precise and prompt identification is needed for effective monitoring of phytophagous Thysanoptera as quarantine pests, to prevent dispensable crop damage. Semi-automated identification using an ANN model based on morphological data could offer a way for easier monitoring and an alternative to e.g. molecular diagnostic method. Optimal architecture (16,4,2). (Fedor et al. 2009, Fedor et al. subm. 2012).

Conclusions

ANN methodology can be applied for classification of any objects for which suitable characters can be found and specified. ANN represent an efficient tool for fast identification of zoological objects. ANN are able to determine even hardly distinguishable species or to mark quite new species from study area or as sp.n. Higher number of objects for the training of ANN improve the quality of identification. The high success and efficacy of ANN is demonstrating wide applicability of this approach in zoological practice.

Acknowledgements

For financial support the European Social Fund and the Ministry of Education of the Czech Republic (CETPO project CZ.1.07/2.3.00/20.0166) , Masaryk University (Department of Botany and Zoology) and the Slovak Science Foundation VEGA (grant No. 1/0137/11) are acknowledged.