

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

Carlos Fernando Crispim Junior

**USO DE DESCRITORES MORFOLÓGICOS E CINEMÁTICOS
NA IDENTIFICAÇÃO AUTOMÁTICA DE
COMPORTAMENTOS
DE ANIMAIS DE LABORATÓRIO**

Tese submetida ao Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina para a obtenção do Grau de Doutor em Engenharia Elétrica

Orientador: Prof.Dr. José Marino Neto

Co-orientador: Prof.Dr. Fernando Mendes
de Azevedo

Florianópolis

2011

Catálogo na fonte pela Biblioteca Universitária da
Universidade Federal de Santa Catarina

C932u

Crispim Junior, Carlos Fernando

Uso de descritores morfológicos e cinemáticos na identificação automática de comportamentos de animais de laboratório [tese] / Carlos Fernando Crispim Junior ; orientador, José Marino Neto, co-orientador, Fernando Mendes de Azevedo. - Florianópolis, SC, 2011.

137 p.: il., grafs., tabs.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.

Inclui referências

1. Engenharia elétrica. 2. Morfologia (Animais). 3. Cinemática. 4. Animais - Comportamento. 5. Redes neurais (Computação). 6. Bioengenharia. I. Marino Neto, José. II. Azevedo, Fernando Mendes de. III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Elétrica. IV. Título.

CDU 621.3

Carlos Fernando Crispim Junior

**USO DE DESCRITORES MORFOLÓGICOS E CINEMÁTICOS
NA IDENTIFICAÇÃO AUTOMÁTICA DE
COMPORTAMENTOS DE ANIMAIS DE LABORATÓRIO**

Esta Tese foi julgada adequada para obtenção do Título de Doutor em Engenharia Elétrica, e aprovada em sua forma final pelo Programa de Pós-graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

Florianópolis, 09 de setembro de 2011.

Prof. Patrick Kuo-Peng, Dr.

Coordenador do Programa de Pós-graduação em Engenharia Elétrica

Banca Examinadora:

Prof. Dr. José Marino Neto,

Orientador

Universidade Federal de Santa Catarina

Prof. Dr. Fernando Mendes de Azevedo,

Co-Orientador

Universidade Federal de Santa Catarina

Profa. Dra. Cinthia Itiki,

Universidade de São Paulo

Prof. Dr. Renato Garcia Ojeda,
Universidade Federal de Santa Catarina

Prof. Dr. Eros Comunello
Universidade do Vale do Itajaí

Prof. Dr. Jefferson Luiz Brum Marques
Universidade Federal de Santa Catarina

A todos aqueles que contribuíram para
o nascimento dessa fagulha de saber
que habita em meu peito.

AGRADECIMENTOS

Agradeço a Deus por essa jornada gratificante, e pelo apoio nas horas difíceis. A minha família e noiva pela compreensão nas ausências, e principalmente pelo amor e zelo que dedicaram a mim. Esse amor foi essencial para o desenvolvimento deste trabalho.

Ao professor José Marino-Neto pela orientação singular, e os conselhos para o futuro. A Profa. Dra. Cilene pelo apoio no doutorado e em meus planos para o pós-doutorado.

Ao prof. Dr. Fernando Azevedo pela co-orientação, e pela demonstração dos valores que devemos cultivar no decorrer dos anos para construção de um mundo mais justo e sensato.

Ao companheiro de doutorado Dr. Miguel Sovierzoski e a Dra. Daniela Suzuki pelos conselhos técnicos e de vida nas conversas de corredor.

Aos colegas de laboratório Alexandre Hoeller, Tiago de Souza Santos e Fernando Melleu pelo auxílio nos procedimentos experimentais, e pela motivação diária com seu bom humor e irreverência.

Aos amigos Diego Laucsen, Vander Vígolo, Cesar Pamplona, Rafael Briese, Márcio Beppler, Francisco de Assis, Yerko Vargas, Willi Osaka, Marcos Signori, Juliano Leite, Geovani Scolaro, Angeluz Canena, e William Castañeda pelo companheirismo nessa jornada. Devo a vocês parte do meu saber, e muitas das alegrias desses quatro anos de doutorado.

Ao Prof. Dr. Samir Mussa pelas orientações no estágio de docência, e aos docentes do IEB-UFSC pela experiência compartilhada. Aos professores do departamento de Ciências Biológicas e Farmacologia pelo apoio nos testes do software o ETHOWATCHER®, pela adoção dela nos procedimentos experimentais de seus laboratórios.

Ao IEB-UFSC, a CAPES e o CNPq pelo fomento de meu doutorado, e o prêmio “Fundadores IEB-UFSC”.

Não reclame dos ventos, nem espere que eles mudem, ajeite as velas e faça seu próprio caminho.

RESUMO

O comportamento animal é um sinal biológico pouco explorado pelas disciplinas de processamento de sinais e inteligência computacional em Engenharia Biomédica. As neurociências usam registros e quantificações do comportamento animal para examinar os mecanismos neurais de controle comportamental. Estes registros são geralmente realizados por um observador humano, e estão sujeitos a vieses de interpretação (*e.g.*, cansaço, experiência, e ambiguidades entre as categorias). Neste trabalho, examinou-se o uso de descritores de imagens (morfológicos: como área e comprimento; e cinemáticos, como distância percorrida e variação angular) como parâmetros de entrada de redes neurais artificiais (RNAs), na identificação automática de comportamentos de animais de laboratório. Os descritores foram extraídos de comportamentos de ratos Wistar em arenas de campo aberto (locomoção: LOC, imobilidade: IMO, limpeza corporal: LIC, exploração vertical: EXP), tratados com cafeína (2 ou 6 mg/Kg) ou com seu veículo (salina), usando um software de etografia e rastreamento desenvolvido durante essa tese (o ETHOWATCHER®). Empregou-se RNAs perceptron de múltiplas camadas (MLP), avaliadas por múltiplos índices de desempenho de diagnóstico (AUC, Kappa). Os descritores foram previamente avaliados quanto a sua relevância na diferenciação entre os comportamentos usando o teste estatístico de Kruskal-Wallis. Em animais tratados com veículo, as MLPs identificaram $97,3 \pm 2,0$ % dos casos de IMO (AUC, média \pm desvio padrão); $95,6 \pm 8,0$ % de LOC; $94,6 \pm 3,0$ % de EXP; e $83,6 \pm 16,0$ % de LIC. Em animais tratados com cafeína, obteve-se: $85,2 \pm 1,8$ % em IMO; $83,5 \pm 0,9$ % em LOC; $67,0 \pm 2,0$ % em EXP; e $78,0 \pm 10,0$ % em LIC. Os resultados indicam que as MLP usando os descritores cinemáticos e morfológicos identificam com sucesso variável os comportamentos investigados. As diferenças estatisticamente significantes entre os desempenhos dos classificadores usando parâmetros relevantes e aqueles usando irrelevantes validaram o uso do teste Kruskal-Wallis na seleção de descritores adequados para a identificação comportamental. A redução de desempenho da MLP em comportamentos de animais tratados com cafeína em dose sub-efetiva (0,2 mg/Kg) pode sugerir que os procedimentos aqui utilizados são capazes de detectar variações em padrões morfológicos e cinemáticos dos comportamentos (Mann-Whitney, $p < 0,05$) não detectáveis pelos procedimentos usuais de análise comportamental. Embora reduzido, o desempenho da MLP foi superior ao medido em observadores iniciantes no registro comportamental de um rato ingênuo a tratamento (Kappa: 35,48%), evidenciando a viabilidade do uso dessas RNAs na avaliação de alterações em padrões comportamentais.

Palavras-chave: Morfologia, Cinemática, Comportamento Animal, Redes Neurais Artificiais, Bioengenharia

ABSTRACT

Animal behavior is a biological signal poorly explored by signal processing and computational intelligence courses in Biomedical Engineering. Neurosciences use animal behavioral recordings and quantifications to examine the neural mechanics of behavioral control. The behavioral recordings are generally done by a human observer, which is influenced by interpretation biases (*e.g.*, tiredness, experience, and categories ambiguity). We examined the use of image features (morphological: area and length; kinematics: distance travelled and angular variation), as input-parameters in artificial neural networks (ANN) for the automatic identification of behaviors of laboratory animals. Image features were extracted from behaviors of *Wistar* rats in an open field arena (locomotion: LOC, immobility: IMO; grooming: GRO, rearing: REA) treated with caffeine (2 or 6 mg/kg) or vehicle, using a software developed during this thesis for Ethograph and *video-tracking* (the ETHOWATCHER®). We used Multilayer Perceptron (MLP) ANN, evaluated by multiple performance indices (*e.g.*, area under ROC curve – AUC, Kappa). Image features were previously evaluated on their relevance for the differentiation among the behavioral categories using the Kruskal-Wallis test. For animals treated with vehicle, MLPs identified 97.3 ± 2.0 % of the IMO samples (AUC, mean ± standard deviation); 95.6 ± 8.0 % of the LOC, 94.6 ± 3.0 % of the REA, and 83.6% ± 16.0 % of the GRO. For animals treated with caffeine, MLPs identified 85.2 ± 1.8 % of the IMO; 83.5 ± 0.9 % of the LOC; 67.0 ± 2.0 % of the EXP; and 78.0 ± 10.0 % of the LIC. These results showed that the MLP-ANN using kinematic and morphological features identified the investigated behaviors with a variable success. The statistical differences in performance among classifiers trained with relevant features and those trained with irrelevant features had validated the use of Kruskal-Wallis test to select appropriate features for behavioral identification. MLP performance reduction in animals treated with a sub-effective dose of caffeine (0.2 mg/kg), suggested that the procedures herein used are capable of detect behavioral morphological and kinematics patterns changes (Mann-Whitney, $p < 0.05$), undetectable by the usual procedures of behavioral analysis. Even reduced, the MLP performance were superior to the results of beginner human observers in the behavioral recordings of a rat treated with vehicle (Kappa: 35.48%), showing the viability of ANNs use for the evaluation of changes in behavioral patterns.

Keywords: Morphology, Kinematics, Animal behavior, Artificial Neural Networks, Bioengineering

LISTA DE FIGURAS

Figura 1. Sequência de posturas que exemplificam o comportamento de exploração vertical.....	28
Figura 2. Imagem extraída de uma posição intermediária de um evento de locomoção. Vista isoladamente a imagem pouco difere de uma imagem de um animal imóvel.....	29
Figura 3. Gaivotas expressando comportamentos sociais.....	34
Figura 4. Modelos animais usados no estudo da emoção e ansiedade: Campo Aberto (A) e Labirinto Elevado em Cruz (B) [35].....	34
Figura 5. Efeitos do observador registrando o comportamento de um animal em campo.....	36
Figura 6. Imagens do rastreamento da atividade de dois ratos na arena de campo aberto. A imagem A ilustra o animal ingênuo ao tratamento, enquanto a imagem B ilustra o animal tratado com uma droga estimulante (caféina).....	39
Figura 7. ACTIMOTIL (A) e PAS (B).....	40
Figura 8. Estrutura do LABORAS TM e exemplos de padrões de vibração dos comportamentos.....	41
Figura 9. Arquitetura de um Sistema de Visão Computacional.....	43
Figura 10. Estrutura do <i>Perceptron</i> de Múltiplas Camadas alimentado adiante	49
Figura 11. Curvas do erro médio quadrático de treinamento e de validação de uma RNA.....	52
Figura 12. Comparação de desempenho de dois classificadores através de suas curvas ROC. Os pontos P1, P2 e P3 representam o desempenho de diagnóstico (em pares de sensibilidade e especificidade) de diferentes pontos de corte da função de ativação do neurônio de saída de classificadores de resposta binária [68][66].....	54
Figura 13. Locomoção a cada 10 quadros.....	62
Figura 14. Imobilidade a cada 10 quadros.....	62
Figura 15. Limpeza corporal a cada 10 quadros.....	63
Figura 16. Exploração Vertical a cada 10 quadros.....	63
Figura 17. Imagem de um rato no ambiente de campo aberto (A), seguida da imagem resultante do procedimento de subtração, com ilustração do ponto estimado como centro da área do animal (B).....	70
Figura 18. Método de Estimação do Comprimento do Animal. A Figura A ilustra o cálculo de uma reta i dado o ângulo α . A Figura B ilustra o conjunto de retas produzido pelo método para um intervalo de ângulos de -89 a 90	71
Figura 19. A área do animal é estimada como a quantidade de pixels com intensidades de cor diferente do plano de fundo na imagem resultante da segmentação (A). O número de pixels alterados é a quantidade de pixels que mudaram de intensidade de cor na comparação entre as matrizes pós procedimento de segmentação referentes a imagem (frame) atual (n) e da anterior ($n-1$). A figura B ilustra essa comparação sobrepondo a informação pré-segmentação de dois quadros consecutivos.....	72

Figura 20. Representação da Informação temporal dos quadros de um comportamento	74
Figura 21. Exemplos de funções logísticas.....	76
Figura 22. Histogramas dos atributos de assimetria da área e de médias geométrica e harmônica do comprimento do animal	80
Figura 23. Atributos de média e desvio padrão dos descritores comportamentais (distância percorrida, número de pixels alterados, e comprimento do animal) por categoria comportamental. (*) indica diferenças estatisticamente significantes entre a locomoção e as outras categorias analisadas (Kruskal-Wallis, $p < 0.05$).	84
Figura 24. Histogramas do número de quadros das amostras comportamentais das categorias comportamentais de locomoção (A), imobilidade (B), exploração vertical (C), e limpeza corporal (D).....	87
Figura 25. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de Locomoção, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney).	90
Figura 26. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de Imobilidade, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney) 90	90
Figura 27. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de Limpeza Corporal, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney).	91
Figura 28. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de exploração vertical, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney).	92
Figura 29. Desempenho de detecção dos classificadores em amostras de 1 segundo (30 quadros) dos eventos comportamentais. O eixo vertical apresenta a resposta do classificador (1 – comportamento detectado, 0 – outro evento). O eixo horizontal apresenta a sequência de categorias de comportamentos usados no teste do classificador (L – Locomoção, I – Imobilidade, E – Exploração Vertical, A – Limpeza Corporal). Para os comportamentos de exploração vertical e imobilidade são ilustrados dois classificadores, pois os índices AUC e Kappa discordaram no classificador de maior desempenho de detecção.....	97
Figura 30. Desempenho de identificação comportamental dos classificadores nas 10-dobras da validação cruzada segundo diferentes índices de desempenho, por grupo de atributos usado no treinamento.....	100
Figura 31. Comparação de diferentes índices de desempenho na avaliação da capacidade de identificação de comportamentos de limpeza corporal por	

classificadores treinados a partir de atributos não selecionados pela prova estatística (segundo os índices AUC, eficiência, Kappa, sensibilidade e especificidade).....	100
Figura 32. Tela Inicial – Tutorial.....	113
Figura 33. Tela de Calibração do Instrumento.....	115
Figura 34. Tela de Criação de Catálogo.....	116
Figura 35. Tela de Etografia.....	117
Figura 36. Exemplo de Relatório de Etografia.....	118
Figura 37. Tela de Análise de Atividade.....	119
Figura 38. Exemplo de Relatório de Atividade.....	120
Figura 39. Imagens do rastreamento da atividade de dois ratos na arena de campo aberto. A imagem A ilustra o animal ingênuo ao tratamento, enquanto a imagem B ilustra o animal tratado com uma droga estimulante (caféina)	120
Figura 40. Referências dos pontos da arena de campo aberto nos quais foi medido o iluminamento	122
Figura 41. Ambiente de teste do Movimento Circular.....	125
Figura 42. Rastreamento do objeto nas três velocidades testadas	125
Figura 43. Sobreposição das imagens resultantes do procedimento de segmentação do objeto rastreado no período de uma volta, em cada uma das velocidades testadas (respectivamente: 69,4 cm/s; 153,36 cm/s; e 241,69 cm/s).	127
Figura 44. Procedimento de Teste de Movimento Circular com um Retângulo	128

LISTA DE QUADROS

Quadro 1. Técnicas de segmentação usadas na análise do comportamento animal	44
Quadro 2. Quadrado latino da sequencia de administração das drogas em cada animal	61
Quadro 3. Descritores de imagem analisados	69
Quadro 4. Atributos estatísticos de cada descritor morfológico/ cinemáticos selecionados pela prova estatística para a diferenciação entre as categorias comportamentais.....	82
Quadro 5. Grupo de Controle II - Complemento da Pré-seleção	85

LISTA DE TABELAS

Tabela 1. Frequência dos comportamentos por grupo de tratamento.....	64
Tabela 2. Número de Neurônios nas Camadas das RNA-MLP de maior desempenho	76
Tabela 3. Parâmetros usados no algoritmo de treinamento dos MLP	77
Tabela 4. Nível de desempenho de identificação dos comportamentos segundo índices de eficiência, AUC, e Kappa	93
Tabela 5. Desempenho dos classificadores na validação cruzada de 10 dobras usando amostras de comportamentos extraídas de vídeos de animais tratados com cafeína.....	94
Tabela 6. Nível de significância das diferenças estatísticas de desempenho entre a identificação de comportamentos de animais tratados com cafeína 6mg/kg em relação à identificação de amostras de comportamentos de animais sem tratamento	95
Tabela 7. Nível de significância das diferenças estatísticas de desempenho entre os classificadores usando grupos de atributos pré-selecionados, e aqueles usando apenas a média e o desvio padrão, na identificação de comportamentos em amostras de ratos tratados com cafeína 2 e 6 mg/kg.	95
Tabela 8. Desempenho dos classificadores na validação cruzada de 10-dobras usando amostras de 30 quadros (\approx 1 segundo).....	96
Tabela 9. Nível de significância das diferenças estatísticas de desempenho entre os classificadores usando atributos pré-selecionados e aqueles usando apenas a média e o desvio padrão dos descritores, na identificação de comportamentos em segmentos de vídeo de 1 segundo de duração.	98
Tabela 10. Nível de significância do teste de Mann-Whitney na avaliação de diferenças estatísticas de desempenho dos classificadores na identificação nos comportamentos do conjunto de validação e nos comportamentos com 1 segundo de duração.....	98
Tabela 11. Distância Percorrida pelo objeto em relação aos Posicionamentos	123
Tabela 12. Comprimento do objeto (pixels) em relação aos Posicionamentos	123
Tabela 13. Ângulo de orientação da reta que caracteriza o comprimento do animal e sua variação entre quadros consecutivos em relação aos posicionamentos dos objetos	123
Tabela 14. Número de Pixels do objeto em relação aos Posicionamentos	124
Tabela 15. Número de Pixels alterados em relação aos Posicionamentos	124

LISTA DE ABREVIATURAS E SIGLAS

ANOVA	- análise de variância
AUC	- area under curve
HMM	- hidden markov model
IEB	- Instituto de Engenharia Biomédica
INPI	- Instituto Nacional de Propriedade Intelectual
IP	- intra-peritoneal
FN	- falso-negativo
FP	- falso-positivo
MLP	- <i>multilayer perceptron</i>
PAS	- Photobeam Activity System
RNA	- redes neurais artificiais
ROC	- <i>receiver operating characteristic</i>
SC	- subcutânea
VN	- verdadeiro-negativo
VP	- verdadeiro-positivo
UFSC	- Universidade Federal de Santa Catarina

SUMÁRIO

SUMÁRIO	25
1. INTRODUÇÃO	27
1.1. JUSTIFICATIVA	30
1.2. OBJETIVOS	31
1.2.1. Objetivos Específicos	31
2. FUNDAMENTAÇÃO TEÓRICA	33
2.1. INTRODUÇÃO	33
2.2. MÉTODOS DE ETOGRAFIA	37
2.3. PROCESSAMENTO DE IMAGENS	42
2.4. DESCRITORES MORFOLÓGICOS E CINEMÁTICOS	45
2.5. REDES NEURAIS ARTIFICIAIS	48
2.6. AVALIAÇÃO DA GENERALIZAÇÃO	50
2.7. AVALIAÇÃO DO DESEMPENHO DE IDENTIFICAÇÃO	52
2.8. PRÉ-SELEÇÃO DOS PARÂMETROS DE ENTRADA	57
3. MATERIAIS E MÉTODOS	59
3.1. PROTOCOLO EXPERIMENTAL	59
3.1.1. Ambiente Experimental	59
3.1.2. Animais	60
3.1.3. Soluções e doses administradas	60
3.1.4. Administração das Drogas	61
3.1.5. Catálogo de Comportamentos	62
3.2. ETHOWATCHER®	64
3.3. DESCRITORES COMPORTAMENTAIS	68
3.4. ATRIBUTOS DE ESTATÍSTICA DESCRITIVA	72
3.5. TREINAMENTO E AVALIAÇÃO DAS REDES NEURAIS ARTIFICIAIS	75
3.6. VALIDAÇÃO DA PRÉ-SELEÇÃO DOS DESCRITORES COMPORTAMENTAIS	78
3.6.1. Grupo de atributos pré-selecionados	81

3.6.2.	Grupos de atributos de controle	83
3.7.	AVALIAÇÃO DOS DESCRITORES	85
4.	RESULTADOS	89
4.1.	INTRODUÇÃO	89
4.2.	PRÉ-SELEÇÃO DE PARÂMETROS	89
4.3.	IDENTIFICAÇÃO DOS COMPORTAMENTOS	92
4.4.	AVALIAÇÃO DOS ÍNDICES DE DESEMPENHO	98
5.	DISCUSSÃO	101
6.	CONCLUSÕES.....	109
7.	DESENVOLVIMENTOS FUTUROS.....	111
	APÊNDICE A – ETHOWATCHER®	113
	APÊNDICE B – VALIDAÇÃO DOS DESCRITORES.....	121
	TESTE 01 – ESTABILIDADE DOS DESCRITORES EM OBJETOS PARADOS	122
	TESTE 02 – DISTÂNCIA PERCORRIDA REAL E ESPERADA	124
	TESTE 03 – ORIENTAÇÃO DO OBJETO E VARIAÇÃO ANGULAR ...	127
	CONCLUSÕES.....	128
	REFERÊNCIAS.....	129

1. INTRODUÇÃO

O comportamento animal é um sinal biológico pouco explorado pelas disciplinas de engenharia de processamento de sinais. As disciplinas da área de Neurociências (neurofisiologia, psicologia, farmacologia) o estudam para examinar os mecanismos de ação e os efeitos de drogas sobre o organismo (*e.g.*, ingestão de alimentos ou mecanismo de defesa), e para descobrir e compreender os processos neuroquímicos e/ou neurofisiológicos subjacentes a ele [1][2][3][4] *apud* [5][6] *apud* [5][7]. O termo “comportamento” pode abarcar qualquer tipo de atividade que acarrete mudança na forma ou função de uma parte (ou de todo) corpo de um animal, em relação ao que havia anteriormente. Cada atividade pode constituir uma categoria descritiva [8].

O registro automatizado do comportamento pode ser feito (para um conjunto limitado de propósitos) de forma indireta, através da medição dos sinais produzidos pelas ações do animal no ambiente experimental. Acoplam-se sensores a arena experimental (*e.g.*, de vibração, de infravermelho, de pressão), e associam-se intervalos dos valores de resposta desses sensores as categorias comportamentais a detectar [3][9][10][11][12][13]. Por exemplo, durante o comportamento de exploração vertical de um rato, o animal fica apoiado sobre suas patas traseiras, e escora as patas dianteiras na parede da arena, tendo acesso a uma área superior do ambiente experimental (Figura 1). Sistemas automatizados acoplam sensores de infravermelho em uma altura da parede da arena, a qual o animal somente tem acesso durante a exploração vertical (Figura 1, postura 5). Cada interrupção do referido sensor é contabilizada como uma ocorrência da categoria. No entanto, algumas categorias comportamentais (*e.g.*, limpeza corporal, comer) não produzem efeitos relevantes nos sensores como o estereótipo característico da exploração vertical [14]. São comportamentos em que ocorrem alterações sucintas em partes específicas do corpo do animal, de difícil descrição para um algoritmo computacional (*e.g.*, limpeza corporal: o animal realiza movimentos de limpeza sobre sua cabeça e patas).

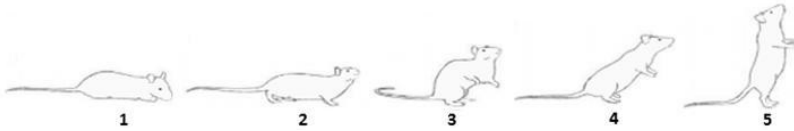


Figura 1. Sequência de posturas que exemplificam o comportamento de exploração vertical

Fonte: adaptado de BENJAMINI *et al.* [15]

Em resumo, ainda que o registro do comportamento feito por um observador humano esteja sujeito a uma variada gama de vieses (*e.g.*, inconsistências devido a níveis distintos de experiência, cansaço, estresse), não há um sistema automatizado capaz de reproduzir suas habilidades de registro comportamental. O observador humano é capaz de registrar o comportamento apenas com a informação visual de sua execução, e em geral o faz independente do ambiente em que o animal se encontre. Estudos envolvendo processamento de imagens tentam identificar quais padrões morfológicos são usados pelo observador (*e.g.*, área do animal, comprimento, padrões de contorno, padrões espectrais) [16][17][18], com intuito de aprimorar o procedimento de identificação comportamental, e detectar comportamentos mais sutis. Contudo, tem-se discutido que nem toda informação visual usada pelo observador é de natureza morfológica. Estudos recentes indicam que em alguns comportamentos os padrões cinemáticos das ações do animal são mais importantes (e menos ambíguos) que as alterações morfológicas da área de seu corpo [15]. Por exemplo, imagens intermediárias de uma locomoção vistas isoladamente são semelhantes (se não iguais) às de imobilidade (vide Figura 2).



Figura 2. Imagem extraída de uma posição intermediária de um evento de locomoção. Vista isoladamente a imagem pouco difere de uma imagem de um animal imóvel.

Além disso, na maioria das categorias o registro de eventos comportamentais refere-se a um período de tempo (conjunto de imagens). É nesse período de tempo que o animal apresenta uma determinada sequência de posturas (morfologicamente estereotipadas), desloca-se de um ponto a outro na arena, ou permanece estático em uma parte da arena. Ou seja, tenha um perfil cinemático ou morfológico, é a variação ou invariância desses perfis durante o período que constitui o evento que descreve a categoria comportamental. Como os descritores comportamentais geralmente se referem a informação de uma imagem (ou no máximo a um par delas), é necessário usar de métodos que representem o relacionamento de cada descritor durante o conjunto de imagens de um evento. Por exemplo: descritores morfológicos combinados com heurísticas [16], ou modelos de inteligência computacional que considerem o relacionamento existente entre as imagens contidas em um evento (*e.g.*, memória dos dados) [17] [19][20][21].

O estudo de descritores de imagens e máquinas de aprendizagem deve lidar com questões como a alta dimensionalidade de parâmetros de entrada. A adição indiscriminada de parâmetros de entrada a modelos computacionais pode deixar de agir como um reforço, e atuar como uma fonte de ruído para o processo de inclusão de conhecimento [17][22] [23][24]. É importante desenvolver procedimentos de seleção que evidenciem (segundo um ou mais índices) aqueles descritores potencialmente relevantes para descrição dos eventos de interesse (*e.g.*,

categorias comportamentais). O uso de um grupo seletivo de descritores (previamente avaliados quanto a sua relevância) provavelmente estará menos suscetível a outras tendências presentes nos dados que não aquelas a mapear (*e.g.*, ruídos) [19].

1.1. JUSTIFICATIVA

A identificação de comportamentos em vídeos experimentais, usando procedimentos de processamento de imagens e inteligência computacional apresenta resultados promissores, ainda que no presente não sejam aplicáveis na rotina do procedimento experimental. Parte dos exemplos é capaz de reconhecer apenas uma categoria comportamental [18][21]. Em sistemas de detecção que sejam sensíveis a um conjunto mais amplo de comportamentos [17], os resultados indicam a identificação correta de comportamentos visualmente similares (desde que morfologicamente estereotipados), porém ainda falham em identificar categorias distintas (*e.g.*, locomoção e imobilidade), provavelmente pela ausência de métodos de descrição da componente cinemática desses comportamentos [15]. Métodos recentemente desenvolvidos concentram-se em discriminar os efeitos de diferentes tratamentos (*e.g.*, substâncias geradoras de ansiedade) sobre os padrões gerais do comportamento animal, devido às dificuldades existentes em caracterizar os eventos precisamente [16] [25] [26]. Em resumo, esses métodos verificam se a droga em estudo alterou o padrão geral de atividade dos animais, sem identificar quais comportamentos foram alterados, nem reconhecer como esses eventos foram modificados.

Parte das falhas dos sistemas de registro/identificação automática de comportamentos deve-se a dificuldades no treinamento do modelo de rede neural escolhido [19]. A ausência de uso de um índice ou procedimento para avaliação do ponto ótimo para parada de treinamento do modelo tende a resultar em um excesso de treinamento (ou de ajuste) desses modelos computacionais, que se ajustam às características dos exemplos de treinamento. Esse excesso de ajuste torna às redes incapazes de reconhecer outros exemplos das categorias, além daqueles usados durante o treinamento. O exame de índices que auxiliem a determinação do ponto mais adequado para a interrupção do treinamento, além da avaliação do desempenho desses modelos de classificadores foi um dos objetivos do presente projeto.

Além disso, parece importante que os parâmetros de entrada dessas redes também sejam avaliados e selecionados quanto a sua

relevância para a descrição dos eventos de interesse (antes de seu emprego como parâmetros de entrada das redes neurais). Um de nossos objetivos foi o estudo de provas estatísticas para pré-seleção dos parâmetros de entrada (descritores comportamentais) de modelos de RNA empregadas na detecção comportamental.

Neste trabalho foram estudados descritores da cinemática e da morfologia do animal para a representação das características de categorias comportamentais (*e.g.*, locomoção, imobilidade, limpeza corporal) de animais de laboratório. O termo morfologia advém da biologia, e trata-se da descrição da forma e da estruturas de animais e plantas [27]. A cinemática trata do estudo e da descrição do movimento de um objeto, e em geral resume o objeto a uma única partícula, desprezando as influências da massa ou das forças que atuam sobre esse objeto (*e.g.*, distância percorrida, velocidade média, aceleração) [28]. Foram empregados atributos estatísticos para representar a variação desses descritores durante as imagens contidas no período de tempo de cada amostra comportamental. Esses atributos de descritores foram então empregados como parâmetros de entrada de redes neurais artificiais voltadas à identificação comportamental. O teste de Kruskal-Wallis (prova estatística não paramétrica) foi usado e validado como método de seleção de parâmetros de entrada relevantes para a descrição de eventos comportamentais.

1.2. OBJETIVOS

Estudar o uso de descritores comportamentais (cinemáticos e morfológicos) e de seus atributos estatísticos, extraídos de segmentos de vídeo obtidos em situação experimental real, no reconhecimento automático de comportamentos de animais de laboratório, por meio de redes neurais artificiais avaliadas por diferentes índices de desempenho de diagnóstico e de avaliação da concordância de registro comportamental.

1.2.1. Objetivos Específicos

1. Desenvolver uma ferramenta (o ETHWATCHER®) que permita o registro de comportamentos através do teclado do computador (a partir de vídeos digitais), e paralelamente extraia os descritores comportamentais dos segmentos de

- comportamento registrados. Esta ferramenta, bem como a relação dos descritores por ela produzidos, será validada através da extração dos descritores em situações de referência (com resposta conhecida), e com dados de experimentos nos quais ratos estão sob efeito de tratamentos que afetam seus comportamentos;
2. Examinar o uso de procedimentos estatísticos na seleção de descritores comportamentais e atributos estatísticos potencialmente significativos e relevantes como parâmetros de entrada de modelos de RNA destinadas ao reconhecimento de comportamentos, além de investigar a relevância dos descritores selecionados em situações experimentais que alteram o comportamento (e.g., em animais tratados com substâncias neuroativas que afetam o comportamento de locomoção);
 3. Examinar modelos de RNA existentes para construção de classificadores destinados a reconhecer comportamentos de ratos (e.g., locomoção, imobilidade, exploração, limpeza corporal) a partir dos descritores selecionados (2º objetivo); e
 4. Avaliar o desempenho dos modelos de classificadores neurais construídos através de diferentes índices de desempenho de diagnóstico e de concordância de observadores, quanto à eficácia na identificação de comportamentos, e sua capacidade de reconhecimento de exemplos que não estavam presentes na etapa de treinamento (generalização).

2. FUNDAMENTAÇÃO TEÓRICA

2.1. INTRODUÇÃO

O comportamento é um conjunto de respostas (*e.g.*, ações) de organismos vivos (individualmente ou em grupo) em relação a estímulos internos ou externos [1]. Em geral, essa definição abarca qualquer tipo de atividade que acarrete mudança na forma ou função, de uma parte (ou de todo) corpo de um animal, desde que essas mudanças possam ser registradas e analisadas de forma objetiva e quantitativa [7][8].

A Etologia é a ciência que estuda os comportamentos de animais com ênfase nas funções e na evolução histórica de seus padrões de comportamento sob o ponto de vista biológico. Ela avalia aspectos causais, funcionais, ontogenéticos (ligados ao desenvolvimento do indivíduo) e filogenéticos (relacionados à evolução da espécie) do comportamento (*e.g.*, o comportamento social de gaivotas, Figura 3) [4][6] *apud* [2][5][29]. Disciplinas da área de Neurociências (neurofisiologia, psicologia experimental e neurofarmacologia) preocupam-se em investigar os mecanismos de ação dos comportamentos, e os diferentes substratos neurobiológicos envolvidos em sua expressão [3][13]. Esses procedimentos experimentais comparam as alterações comportamentais provocadas por um método em estudo em relação à expressão “padrão” do comportamento (*e.g.*, efeitos de drogas, lesões de circuitos cerebrais, testes comportamentais, e manipulação ou depleção de genes) [30].

Por testes comportamentais entende-se a exposição do animal a um novo ambiente, conhecido por provocar um determinado padrão de resposta comportamental. Por exemplo, o confinamento do animal em um ambiente de paredes altas, que não permita que ele escape (*e.g.*, campo aberto, Figura 3A), tende a provocar comportamentos ligados à ansiedade (animal caminha próximo as paredes, e aumenta a frequência de análises de risco e explorações verticais). Estudos farmacológicos comumente associam um teste comportamental ao estudo de uma droga [30]. Por exemplo, estudos sobre os efeitos de drogas ansiolíticas (supressoras de ansiedade), usam a arena de labirinto em cruz (modelo gerador de ansiedade, Figura 3B) para examinar a efetividade das drogas [31].

O teste de campo aberto é empregado no estudo de drogas que alterem a distribuição espacial da atividade do animal (relacionada com

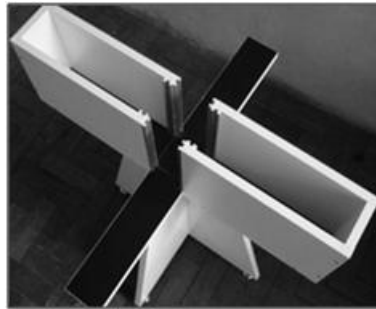
as emoções e a exploração) [32] [33] *apud* [34]. Ele é projetado com uma arena de formato circular (*e.g.*, cilindro) ou quadrado (*e.g.*, caixa de madeira ou plástico), extenso o suficiente para que o animal mova-se livremente, mas com paredes altas que o impeçam de visualizar o exterior. É usado para espécies, como: ratos, camundongos, porcos, coelhos, lagostas, entre outros. Além da locomoção, são observados comportamentos como auto-limpeza ou *grooming* (aqui referido como limpeza corporal), movimentos estereotipados, contatos com novidades, alimentação, exploração vertical (*rearing*) e análise de risco (*risk assessment*)[34].



Figura 3. Gaivotas expressando comportamentos sociais
Fonte: Foto por José Marino Neto, 2008



A



B

Figura 4. Modelos animais usados no estudo da emoção e ansiedade: Campo Aberto (A) e Labirinto Elevado em Cruz (B) [35]

A validação de uma hipótese geralmente fundamenta-se na avaliação do conjunto de categorias comportamentais de interesse para o experimento, quanto à existência de alterações nos padrões de frequência, duração e latência dessas categorias, quando comparados os grupos de animais sob influência do procedimento em teste, e os animais ingênuos ao mesmo. A existência de diferenças significantes entre esses atributos estatísticos é que permite aceitar ou rejeitar a hipótese em teste.

Por exemplo: avaliou-se a hipótese de que os neuropeptídios Orexinas (reguladores do apetite em mamíferos) também tem este efeito em aves [37]. Projetou-se um protocolo experimental com intuito de avaliar o efeito desse tratamento em pombos. As alterações provocadas por esses neuropeptídeos foram avaliadas com base no registro dos comportamentos de comer, beber, dormir, e locomoção. Após a execução do protocolo experimental foram analisados os vídeos de cada experimento, e registraram-se os comportamentos expressos pelos animais. Quantificaram-se os atributos de frequência, duração e latência de cada uma das categorias de comportamentos, e avaliou-se a existência de alterações nesses atributos através da prova estatística ANOVA (Análise de Variância). Ainda que dois animais apresentassem diferenças estatisticamente significantes na ingestão de alimentos, essas diferenças não foram significativas para a prova da hipótese em análise. Portanto, concluiu-se que esses neuropeptídios não têm efeitos na regulação da ingestão de água e alimentos, e nos padrões de frequência e duração dos comportamentos de comer e beber de pombos adultos.

A etapa de registro comportamental (também conhecida como Etografia) consiste na transcrição dos fenômenos comportamentais realizados pelos animais. Essa transcrição deve representar de forma fidedigna os fenômenos comportamentais ocorridos, pois será o cerne da avaliação do aceite/reprovação da hipótese em teste. No entanto, devido à natureza interpretativa do catálogo de comportamentos (instrumento que descreve as categorias a serem registradas) [5], frequentemente ocorrem erros de registro. Entre dois observadores treinados, as discrepâncias entre os registros geralmente referem-se a omissões (*e.g.*, eles concordam sobre a atividade primária do sujeito, porém um observador registra um comportamento 10 segundos antes do outro). Nesses casos as diferenças geralmente anulam-se quando se considera a totalidade dos registros [36, p.86-87].

No entanto, alguns fatores de erro alteram os registros comportamentais. Por exemplo, a falta de informações sobre uma categoria (*e.g.*, devido a mau treinamento ou má redação das definições do catálogo de comportamentos), a natureza ambígua de algumas

categorias comportamentais, erros no uso do método de registro empregado, erros na análise dos dados do registro, erros da perspectiva de aquisição do comportamento, e erros devido à influência da presença do observador sobre o sujeito observado. A Figura 5 ilustra as fontes de erro citadas.

Por exemplo, o estudo do comportamento social de duas espécies de primatas usou de um gravador de som e um relógio. As categorias comportamentais foram representadas por nomes curtos das ações do animal (morder, encarar, auto-limpeza, encaramento, intimidação). Esses nomes indicavam a categoria, o sujeito da ação, e o objeto da ação (e.g., o código “2-1 exsoc” significa que o animal 2 está fazendo exploração social no animal 1). Adicionalmente ao código da ação, registrava-se o tempo de início e fim da ação com base no relógio [36]. Entre os erros de registro encontrados nesse estudo relataram-se a combinação inválida de códigos de categoria, falhas na transcrição textual dos elementos narrados, e defeitos na gravação.

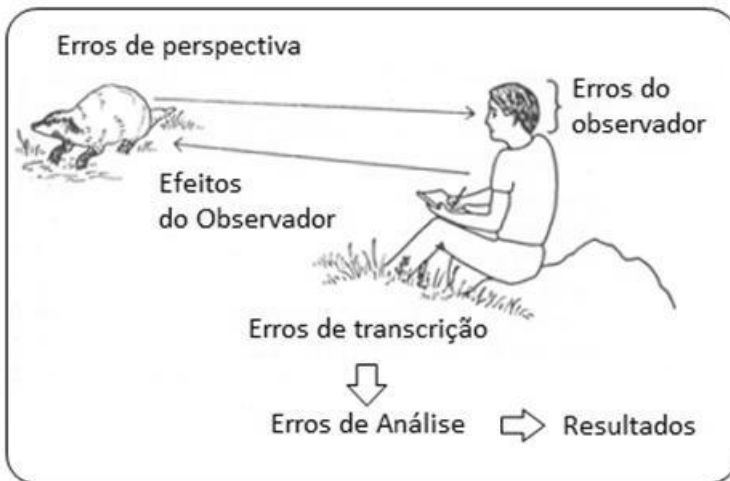


Figura 5. Efeitos do observador registrando o comportamento de um animal em campo.

Fonte: adaptado de LEHNER [5].

Na próxima seção serão descritos alguns instrumentos existentes para o registro comportamental.

2.2. MÉTODOS DE ETOGRAFIA

As técnicas de observação sistemática foram usadas cientificamente pelo menos desde Darwin (1872), no estudo do comportamento do homem e de outros animais. Dentre os métodos usados estavam: as listas de assinalar, a gravação sonora, e os vídeos de filmadoras analógicas. As listas de assinalar eram usadas em casos em que os observadores já conheciam as categorias comportamentais a registrar. A narração (em som ou vídeo) consistia em fazer comentários breves à medida que os comportamentos ocorriam, e eram comumente usados quando ainda se estudava o repertório de comportamentos de uma espécie, ou por pesquisadores inexperientes [36].

Os instrumentos de registro eram (e ainda) são comumente adaptados de forma a evitar alterações na expressão dos comportamentos (efeitos do observador), além de facilitar o registro de comportamentos. Por exemplo, no estudo do comportamento humano, filmadoras analógicas foram modificadas para capturar imagens em uma direção oposta a que estavam apontadas. Essa ilusão permitia aos Etografistas (pesquisadores que registram o comportamento) capturar as expressões naturais do comportamento de pessoas ao seu redor sem que elas percebessem [38].

Com os aperfeiçoamentos tecnológicos surgiram os registradores de eventos. Consistiam de um equipamento dotado de um rolo de papel e um conjunto de penas a tinta ou cera quente, que desenhavam sobre um rolo de papel. O registrador continha um teclado, que continha uma tecla para cada pena do dispositivo. Durante o experimento o observador simplesmente pressionava a tecla previamente associada ao comportamento de interesse, e a respectiva pena deslocava seu traçado em alguns centímetros por alguns instantes, marcando a ocorrência do comportamento. Ao fim do experimento os pesquisadores contabilizavam a frequência e duração de cada comportamento em função dos deslocamentos existentes no traçado das penas [36].

Com o passar dos anos e a popularização dos computadores pessoais, o procedimento de ETOGRAFIA transferiu-se das listas de assinalar e registradores de eventos para os softwares de apoio ao registro comportamental. Nesses softwares o usuário cadastra o catálogo de comportamentos, e com base nesse catálogo, ele transcreve os eventos comportamentais que observa (a partir de um arquivo de vídeo ou a partir da observação do animal em seu ambiente natural em tempo real). Após a etapa de registro, esses softwares automaticamente

contabilizam a frequência, a duração e a latência (tempo decorrido do início do experimento até a primeira ocorrência do mesmo) de cada comportamento (*e.g.*, ETHOLOG [39], ETÓGRAFO [40] – gratuitos). Alguns softwares incluem funcionalidades adicionais, como: a segmentação da sequência original de ocorrência dos comportamentos em intervalos de tempo (*e.g.*, ETÓGRAFO), algoritmos para o estudo da existência de padrões temporais nas sequências comportamentais transcritas, e até a avaliação da concordância intra e inter-observador (*e.g.*, JWATCHER [41], gratuito).

Complementares aos sistemas de apoio ao registro direto do comportamento, surgiram os sistemas de *video-tracking* (análise de atividade em vídeos digitais). Eles permitem ao pesquisador visualizar a distribuição da atividade do animal no ambiente experimental (distribuição espacial). Logo, a cinemática do movimento dos animais e os pequenos comportamentos até então também descartados passaram a ser avaliados [38]. Esses softwares baseiam-se no uso de arquivos de vídeos (geralmente adquiridos previamente por filmadoras, câmeras digitais, ou webcams – de uma perspectiva superior ou lateral da arena comportamental). Técnicas de processamento digital de imagens são empregadas para separar (segmentar) o animal do plano de fundo [42][43], estimar o centro da área de seu corpo, e calcular dados relativos à sua atividade (*e.g.*, velocidade média, velocidade angular absoluta, caminho percorrido, tempo de permanência em regiões de interesse) [44][45][46][47][48].

A combinação de ambas as categorias de ferramentas é usada por pesquisadores para analisar diferentes aspectos da expressão comportamental, e ilustrar de forma qualitativa (através das imagens da atividade animal) e quantitativa (índices de atividade) os resultados de experimentos comportamentais. A Figura 6 ilustra duas imagens da atividade produzida por um animal num período de 10 min. na arena comportamental. Cada imagem representa uma de duas situações distintas: a primeira (A), um animal sem tratamento (A); e a segunda o mesmo animal sob o efeito de cafeína (droga estimulante, B). Por exemplo, os índices obtidos da análise de atividade do animal servem de parâmetro para avaliar diferenças estatísticas entre os padrões comportamentais dessas imagens.

A correlação da informação etográfica (do registro comportamental) com os índices de atividades (*e.g.*, distância percorrida, velocidade, alterações angulares de direção) são fundamentais no estudo detalhado e quantitativo dos padrões

comportamentais. No entanto, são raras as ferramentas que permitem ambas as análises.

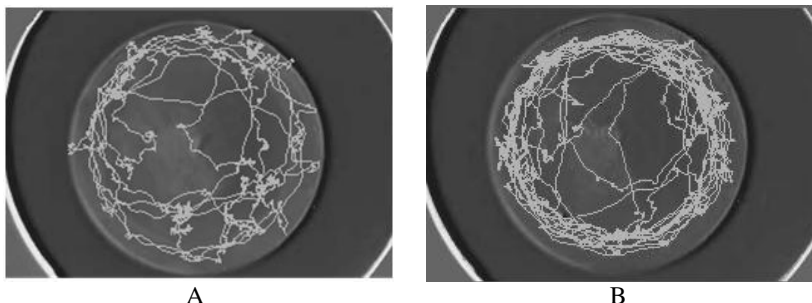


Figura 6. Imagens do rastreamento da atividade de dois ratos na arena de campo aberto. A imagem A ilustra o animal ingênuo ao tratamento, enquanto a imagem B ilustra o animal tratado com uma droga estimulante (cafeína)

Como ambos os tipos de ferramentas são usados pelo observador para quantificar e analisar o fenômeno comportamental, seus resultados estão sujeitos às imprecisões características do observador (*e.g.*, devido a diferenças em experiência, estresse, fadiga, má qualidade das descrições das categorias), que prejudicam a estabilidade dos registros. Existem alguns sistemas de registro automatizado, que medem o comportamento indiretamente, através do acoplamento de sensores a arena comportamental (*e.g.*: acelerômetros, sensores de pressão, de vibração ou de infravermelho). Esses sistemas associam faixas de valores de resposta do sensor à ocorrência de categorias de comportamentos.

Por exemplo, o uso de sensores ópticos localizados na entrada dos braços fechados de um labirinto elevado em cruz (Figura 3B, seção anterior) indicaria a ocorrência de um comportamento de análise de risco. A interrupção de sensores ópticos posicionados a uma determinada altura da arena experimental (a qual o animal tem acesso apenas ficando em pé sobre as patas traseiras) indicaria o comportamento de exploração vertical. No entanto, o uso desses sistemas serve a um conjunto limitado de categorias, que produzam efeitos relevantes na interação com esses sensores. A Figura 7 ilustra dois sistemas existentes que usam de sensores ópticos: o ACTIMOT / MOTIL SYSTEM© (Figura 7.A) [9]; e o PHOTOBAM ACTIVITY SYSTEM - OPEN FIELD (PAS)© [10], ambos compatíveis apenas com a arena de campo aberto (ambos comerciais).



A



B

Figura 7. ACTIMOTIL (A) e PAS (B).

Fonte: Imagens obtidas dos catálogos de seus fabricantes [9][10].

O sistema LABORASTM (*Laboratory Animal Behaviour Observation, Registration e Analysis System* - Figura 8) é um exemplo baseado no uso de sensores de vibração. Ele captura as vibrações mecânicas causadas pelos movimentos do animal na arena através de sensores de vibração posicionados sob a arena. Os comportamentos são identificados de acordo com os padrões de frequência e amplitude característicos de cada categoria comportamental (Figura 8, *e.g.*, padrões de vibração dos comportamentos de comer, beber, exploração vertical, escalada, imobilidade, atividade locomotora e limpeza corporal). Testes farmacológicos verificaram que as variáveis extraídas pelo LABORASTM reproduzem as alterações comportamentais produzidas pelo uso de tratamentos [11][12].

Conforme discutido, as soluções disponíveis (e comerciais) de identificação automática de comportamentos baseiam-se nos dados de sensores ópticos ou mecânicos (*e.g.*, o LABORASTM, o ACTMOT©). Desta forma, sua aplicabilidade está condicionada ao registro de um conjunto restrito de comportamentos que causem alterações significantes na interação com os sensores.

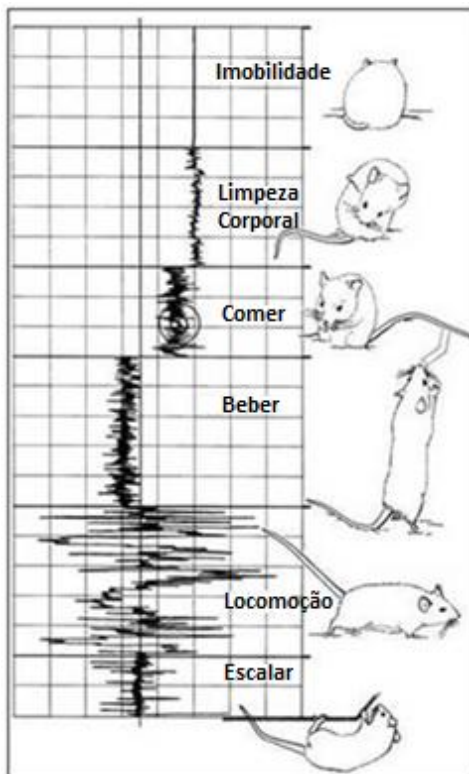


Figura 8. Estrutura do LABORASTM e exemplos de padrões de vibração dos comportamentos.

Fonte: Adaptado de METRIS [49]

Nesta tese foram estudados índices (descritores) que descrevem o comportamental a partir da informação visual contida nas imagens (extraídas de segmentos de vídeos) de animais durante a expressão comportamental. A detecção automática a partir de vídeos experimentais tenta mimetizar a capacidade do observador humano de encontrar e associar evidências visuais únicas a ocorrência das categorias de interesse (*e.g.*, estereótipos morfológicos, como posturas, ou cinemáticos, como o deslocamento da locomoção) [5][8][29]. Uma das vantagens dessa abordagem está em avaliar diretamente o fenômeno comportamental, e não sua consequência no ambiente, como os sistemas baseados na resposta de sensores (*e.g.*, ópticos, de vibração) instalados na arena experimental.

A principal vantagem de ferramentas que identifiquem automaticamente o comportamento reside na estabilidade de suas definições comportamentais ao longo do tempo. Observadores humanos têm dificuldade em repetir os resultados de seus próprios registros (devido às imprecisões e às tendências do observador, *e.g.*, cansaço, estresse, pré-conceitos sobre o experimento), o que dificulta a replicação e comparação de resultados experimentais inter-laboratórios.

Na próxima seção discutem-se os métodos de processamento de imagens envolvidos em sistemas de análise de atividade animal e estudos de detecção de padrões comportamentais.

2.3. PROCESSAMENTO DE IMAGENS

O desenvolvimento de sistemas de visão computacional que mimetizam a habilidade do cérebro humano de reconhecer um objeto de interesse e descrevê-lo, geralmente é decomposto em uma cadeia de procedimentos, cada na qual cada procedimento tem uma função específica (*e.g.*, adquirir a imagem, extrair o objeto da imagem, descrevê-lo). A Figura 9 ilustra um modelo genérico de cadeia de processamento de um sistema de visão computacional, projetado para detectar/reconhecer padrões em imagens.

A cadeia apresentada inicia com procedimentos de baixo nível, que correspondem à aquisição do sinal visual (*e.g.*, através de webcam, filmadora, câmera digital), e de rotinas de pré-processamento (*e.g.*, conversão em uma escala de cores de interesse, realce de uma região, remoção de ruído). O nível intermediário é responsável por extrair o(s) objeto(s) de interesse na etapa de segmentação, e a partir do produto dessa etapa são executados procedimentos de representação e descrição no objeto extraído. Neste trabalho a etapa de descrição do objeto de interesse usa de índices (*e.g.*, descritores morfológicos e cinemáticos) que quantificam as características do objeto, e o diferenciam de seus semelhantes.

O nível alto compreende métodos de tomada de decisão (*e.g.*, classificadores bayesianos, redes neurais artificiais), que auxiliam a classificar/identificar as imagens usando os descritores obtidos nas etapas anteriores, de forma supervisionada ou autônoma [27].

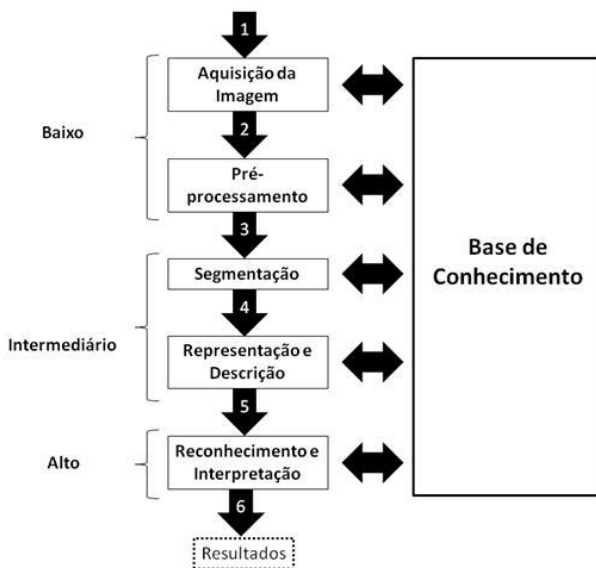


Figura 9. Arquitetura de um Sistema de Visão Computacional
 Fonte: Adaptado de Gonzalez e Woods [27]

As técnicas implementadas em cada etapa da cadeia de processamento dependerão das características do domínio de aplicação do sistema de visão computacional (e.g., rastreamento de animais na cena experimental, de carros em uma avenida; detecção de objetos no caminho de veículos com direção assistida, ou cirurgias guiadas por computador).

O Quadro 1 apresenta técnicas usadas na etapa de segmentação em sistemas de análise da atividade animal (vídeo-tracking), e estudos para a detecção automática de padrões comportamentais. Verifica-se que predomina o uso das técnicas de limiarização e subtração de imagens, complementadas por procedimentos de pós-processamento [16][21][42-51][52][53].

A Limiarização é aplicável quando existe contraste entre o animal e a cena de fundo da imagem de tal forma que é possível selecionar um limiar de cor que separa de forma ótima o animal do plano de fundo da imagem [27]. A subtração de imagens consiste no uso de uma imagem de referência (e.g., ambiente sem a presença do objeto a rastrear), que subtraída da imagem em análise, gera uma imagem contendo apenas o objeto almejado (e.g., animal) [27][43]. Essas duas técnicas são aplicáveis devido a cor do animal desses trabalhos apresentarem

contraste satisfatório em relação a cor do ambiente experimental (cena de fundo), a posição da filmadora de aquisição ser fixa, e a cena de fundo do ambiente ser estática.

Eventuais variações na cor dos pixels do ambiente experimental ocasionadas por variações na fonte de iluminação do cenário ou adição de artefatos ao ambiente durante o vídeo experimental são tratados através de sub-etapas de pós-processamento, como atualização da imagem de referência (subtração de imagens, [43][48][53]), ou o uso de limiares adaptativos (Limiarização, [16][21]).

Autor	Técnica de Segmentação	Ano
HEEREN e COOLS [21]	Limiarização associada a descritores de Fourier	2000
TWINING <i>et al.</i> [52]	Subtração de plano de fundo, e Limiarização, e <i>Active Shape models</i>	2001
SPINK <i>et al</i> [44] (Ethovision)	Subtração de plano de fundo, e Limiarização	2001
LIND <i>et al.</i> [43]	Subtração de plano de fundo, e Limiarização	2005
ZURN <i>et al.</i> [48]	Adaptação da Subtração de plano de fundo	2005
DOLLAR <i>et al.</i> [54]	Fluxo óptico	2005
FROLICH <i>et al.</i> [16]	Limiarização baseada em múltiplos quadros de vídeo	2008
WANG <i>et. al.</i> [55]	Algoritmo CamShift no tracking de humanos	2009
PISTORI <i>et al.</i> [56]	Análise de componentes conectados associada ao algoritmo K-means	2010
JHUANG <i>et al.</i> [53]	Adaptação da Subtração de Plano de Fundo	2010

Quadro 1. Técnicas de segmentação usadas na análise do comportamento animal

Para outros cenários, como o rastreamento de múltiplos animais de laboratório por arena, ou em situações nas quais a coloração dos animais se aproxima da coloração da arena comportamental (e.g., animais marrons em arenas cobertas com serragem), ou a cena de fundo se altera ao longo do experimento verificou-se o uso das técnicas de Fluxo Óptico [54], adaptações do algoritmo K-means em animais de laboratório, e do algoritmo CamShift no rastreamento de humanos em [55].

Como o foco deste trabalho foram as etapas de representação e descrição, e reconhecimento e classificação dos animais, optou-se por usar uma adaptação da técnica de subtração de imagens, complementada por Limiarização. Essa escolha fundamenta-se na existência de contraste satisfatório entre a cor do animal e a cor do ambiente experimental, e a posição da câmera de aquisição ser fixa, como nos exemplos de citados no Quadro 1.

Sugerem-se como desenvolvimentos futuros a comparação de outras técnicas de segmentação (CamShift, Fluxo Óptico e algoritmo K-means) para o aperfeiçoamento do objeto extraído e a análise de sua atividade. Na próxima seção são discutidos os métodos de descrição de imagens vistos em estudos sobre formas de analisar padrões de sinais comportamentais.

2.4. DESCRITORES MORFOLÓGICOS E CINEMÁTICOS

O catálogo de comportamentos descreve o conjunto de comportamentos que deve ser registrado pelo observador. Para cada categoria comportamental são realizadas descrições textuais, geralmente associadas a imagens. Essas imagens ilustram as alterações morfológicas (*e.g.*, posturas, padrões de movimento) características de cada categoria, e tem o intuito de esclarecer as ambiguidades existentes nas descrições textuais.

Sistemas de registro comportamental baseados em vídeos digitais geralmente usam de índices que descrevam a morfologia da área do animal (*e.g.*, a distância entre o focinho e o centróide estimado de sua área, e do centróide a posição estimada de sua cauda). No entanto, índices morfológicos retratam comportamentos que produzam efeitos relevantes na área do animal, como comportamentos estereotipados (*e.g.*, exploração vertical).

A descrição da morfologia das posturas do animal apresentou desempenho de diagnóstico promissor usando grupos de RNA do tipo *perceptron* de múltiplas camadas (comportamento de sentar: Kappa de 83%, Kappa médio de todos os comportamentos de 63%, veja a definição do coeficiente Kappa na seção 2.7). Cada padrão de entrada era composto da informação dos descritores de três quadros de vídeo consecutivos. No entanto, comportamentos visualmente diferentes (*e.g.*, caminhar e sentar) eram confundidos pelo classificador, enquanto comportamentos semelhantes de exploração, provavelmente

morfologicamente distintos, eram detectados. A ausência de uso de procedimentos de avaliação da interrupção do treinamento das redes (e.g., a validação cruzada), somado às diferenças em desempenho nos conjuntos de treinamento e validação, indicam classificadores excessivamente ajustados aos padrões de treinamento, fato que prejudica o desempenho dos classificadores [17].

O uso de descritores do contorno da área de animais (ratos) foi investigado associado a Modelos Ocultos de Markov, na predição de comportamentos de exploração vertical, exploração horizontal e limpeza corporal. A taxa de acerto média foi de 70%. Não foram relatados procedimentos para o tratamento do relacionamento entre as várias imagens (e respectivos descritores) de uma amostra [50].

O contorno do animal descrito através das componentes espectrais da imagem (descritores de Fourier) apresentou uma taxa de acerto médio de 90% na identificação comportamental da limpeza corporal. Os descritores de Fourier eram usados para descrever o contorno do animal, e em seguida seus resultados serviam como parâmetros de entrada de uma RNA do tipo Kohonen (de aprendizagem não-supervisionada). A RNA agrupava as amostras de comportamentos em grupos semelhantes, e as amostras eram rotuladas conforme informações de um banco de dados. Apenas dois ratos foram usados para o treino e teste do procedimento, logo mais estudos devem ser usados para verificar as influências que a variabilidade dos animais tem sobre o resultado do método [21].

Estudos recentes têm privilegiado índices para identificação (ou diferenciação) dos padrões de comportamentos modificados por tratamentos com drogas, ou índices que retratem as diferenças na expressão comportamental em longo prazo devido aos efeitos de doenças degenerativas. Essa abordagem não identifica o evento comportamental propriamente dito, mas as diferenças nos padrões gerais dos descritores de imagem entre os grupos de animais (e.g., animais ingênuos ao tratamento e animais tratados, animais com diferentes tratamentos, animais sadios de doentes) [16][26]. Em um dos casos um descritor do perfil de atividade do animal foi empregado (semelhante ao coeficiente de número de pixels alterados discutido na seção 3.3). Sobre esse descritor foram definidos intervalos de valores (baseados em heurísticas de especialistas) que separavam marginalmente as categorias (e.g., valores <10 foram considerados imobilidade, ≥ 40 locomoção). Ressalta-se o comparativo entre os métodos de classificação *perceptron* de múltiplas camadas e máquinas de vetor de suporte, que apresentaram os maiores desempenhos médios. A avaliação dos classificadores usou

os índices de sensibilidade, especificidade, área sob a curva ROC, e eficiência balanceada. A técnica de validação cruzada foi empregada para evitar o excesso de treinamento [16]. Taxas de acerto de 80% a 90% (índice da Área sob a Curva ROC) suportadas pela medição através dos índices de desempenho relatados, e procedimentos de parada de treinamento reforçam a relevância desses métodos no estudo do registro comportamental. Outros comparativos da discriminação dos efeitos de tratamentos podem ser vistos em [25][26].

Em resumo, os descritores de imagens são índices que representam e descrevem um fenômeno visual. Geralmente são calculados durante ou após o processo de segmentação, e procuram extrair informações relevantes (e preferencialmente simples) para uso por um procedimento de reconhecimento e interpretação de padrões (*e.g.*, morfologia do objeto). Descrevem (ou resumem) diferentes aspectos de uma região (*e.g.*, um objeto de interesse) como textura, contorno, topologia, forma, padrões espectrais [27]. Abordagens recentes combinam diferentes categorias de descritores para representar ou extrair a informação necessária para resolução de problemas complexos de detecção [51].

Privilegiou-se o uso de descritores que extraem informações sobre duas vertentes do fenômeno comportamental: a morfologia (alterações da área do animal), e a cinemática (movimento expresso durante o comportamento). A morfologia matemática procura quantificar partes de imagens, relevantes para um procedimento de descrição e/ou reconhecimento (*e.g.*, contorno, área, organização estrutural) [27]. A cinemática trata do estudo e descrição do movimento de um objeto. Em geral considera o objeto como uma única partícula, desprezando as influências da massa ou das forças que atuam sobre ele. Em resumo, os descritores cinemáticos traduzem as características do movimento do objeto (*e.g.*, distância percorrida, velocidade média, aceleração) [28]. A componente cinemática das categorias de comportamentos tem sido recentemente discutida como parâmetro chave na análise de alterações em padrões comportamentais (*e.g.* usando *softwares* de análise de atividade) [15]. No entanto, o uso de cinemática como padrão de entrada de método de inteligência artificial na detecção pontual de comportamentos é incomum, se não praticamente inexistente [53]. O estudo do emprego dessa classe parâmetros constitui um dos diferenciais desta tese.

Destaca-se a associação de descritores de imagens (na etapa de representação e descrição) com as RNA *perceptrons* de múltiplas camadas (MLP) (na etapa de reconhecimento e interpretação). Os

métodos de aprendizagem mais usados para treinar os *perceptrons* variam do algoritmo de retropropagação de erro de RUMELHART *et al.* [57], ao algoritmo de Levenberg-Marquardt. Além disso, a capacidade desse modelo de incorporar o relacionamento entre os descritores de imagens presentes em uma amostra, através da simples alteração do algoritmo de treinamento adiciona flexibilidade no tratamento da variação dos descritores durante as amostras comportamentais [19][20]. Problemas como o excesso de ajuste aos dados são contornáveis por procedimentos de avaliação do aprendizado (*e.g.*, validação cruzada, parada antecipada) [16][19]. Além disso, parece importante que esses descritores sejam avaliados previamente ao seu emprego como parâmetros de entrada de modelos de RNA.

Os objetivos dessa tese concentraram-se em desenvolver e testar métodos aplicáveis nas etapas de representação e descrição, testados por procedimentos de reconhecimento e interpretação dos dados. A RNA MLP foi escolhida como método de classificação para o teste dos descritores comportamentais produzidos, conforme terceiro objetivo específico. Essa escolha fundamentou-se nos exemplos demonstrados, somado ao fato do laboratório em que essa tese foi desenvolvida possuir experiência no desenvolvimento desse modelo de RNA. As próximas seções desse capítulo abordarão o modelo de RNA MLP, e procedimentos para o treinamento e avaliação desses modelos.

2.5. REDES NEURAIS ARTIFICIAIS

As redes neurais artificiais tentam mimetizar a maneira como o cérebro humano realiza uma tarefa (função) de interesse (*e.g.*, classificação, detecção, filtro). Compõem-se de um conjunto de unidades de processamento simples (neurônios), distribuídas de forma paralela. A propriedade fundamental de uma RNA é a sua habilidade de aprender a partir do ambiente (leia-se conjunto de exemplos), através de um processo iterativo de ajustes aplicados aos pesos sinápticos e níveis de polarização (bias). O reconhecimento de padrões realizado por uma RNA é de natureza estatística, com os padrões sendo representados por pontos em um espaço de decisão multidimensional. O espaço de decisão é subdividido em regiões que representam as classes de dados existentes, o algoritmo de treinamento empregado tem por função ajustar as fronteiras entre essas regiões que separam adequadamente as classes de dados [19][27][58].

Os *perceptrons* de múltiplas camadas alimentados adiante (*feedforward*) estruturam os neurônios artificiais em um conjunto de camadas. Cada neurônio possui uma conexão direta com cada um dos neurônios da próxima camada, e o sinal de entrada se propaga em apenas um sentido (da camada de entrada, pelas camadas ocultas, até a camada de saída). Não existem conexões entre neurônios de uma mesma camada, nem conexões de retorno de um neurônio com neurônios de camadas anteriores, vide Figura 10. As conexões sinápticas entre neurônios (em referência às conexões dos neurônios biológicos) ponderam a influência do sinal do neurônio emissor no neurônio receptor. Os fatores que ponderam as conexões sinápticas recebem o nome de pesos sinápticos. Cada neurônio atua como um somador das conexões que recebe, e aplica sobre essa soma uma função de ativação não-linear (*e.g.*, função sigmóide, tangente hiperbólica). O uso da camada oculta permite que os modelos MLP aprendam tarefas complexas pela extração progressiva das características mais significativas dos padrões de entrada [19][58][59].

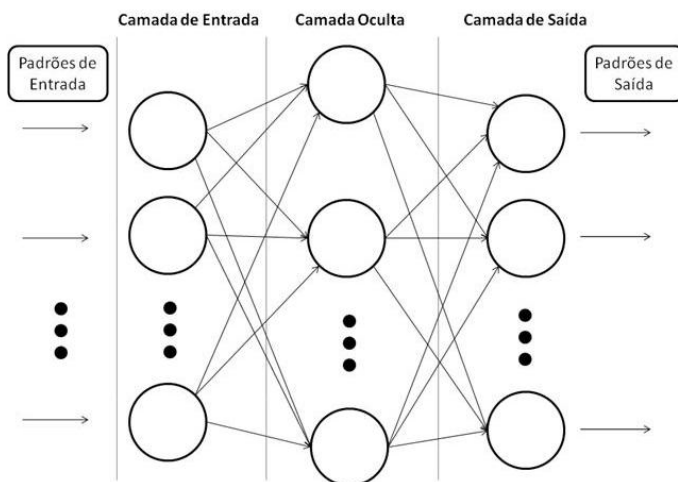


Figura 10. Estrutura do *Perceptron* de Múltiplas Camadas alimentado adiante
Fonte: Adaptado de RUMELHART, HINTON E WILLIAMS [57].

A aquisição de conhecimento (ou treinamento) de um RNA-MLP ocorre por treinamento supervisionado, conhecido como aprendizagem por correção de erro, ou aprendizagem com professor. O professor é uma entidade conhecedora do ambiente em que a RNA deve atuar, e detém as respostas que a rede neural deve apresentar para cada padrão

de entrada fornecido. Em resumo, o treinamento tem por objetivo transferir o conhecimento do professor para a RNA [19].

Um exemplo clássico de algoritmo de treinamento é a retropropagação do erro (*error back-propagation*) com constante de momento, proposto por RUMELHART, HINTON E WILLIAMS [57]. Esse algoritmo subdivide-se em duas etapas: a propagação do sinal e a retropropagação do erro [58]. Na etapa de propagação, um padrão de entrada é inserido na camada de entrada da rede, e calcula-se a saída do sistema para a essa entrada. Compara-se a saída obtida com a saída desejada, e obtém-se o erro associado. A etapa de retropropagação trata da literal propagação do sinal de erro medido da camada de saída para a camada de entrada, com o reajuste dos pesos sinápticos da RNA.

A apresentação consecutiva de todos os padrões de treinamento a rede constitui um período de tempo denominado época. O processo de aprendizagem é repetido por várias épocas até que o ajuste dos pesos sinápticos e os níveis de bias estabilizem, e o erro médio quadrado sobre o conjunto de treinamento convirja para um valor mínimo estabelecido. O treino de uma RNA é uma tarefa longa, que requer várias sessões de treinamento, seguida por uma exaustiva avaliação para selecionar a época de treino que possui os melhores resultados. O erro de treino indica se a RNA está convergindo, porém não é um índice que permite avaliar o aprendizado ou o desempenho de classificação de uma RNA [60]. O estudo de índices que retratem o desempenho de uma época de treinamento e sua capacidade de generalização é foco do quarto objetivo específico deste trabalho, e será discutido na próxima sessão.

2.6. AVALIAÇÃO DA GENERALIZAÇÃO

O treinamento de uma RNA inicia com a seleção de um conjunto de exemplos representativos do domínio de interesse. Esses exemplos de treino serão usados por um algoritmo de treinamento (*e.g.*, retropropagação do erro) para calcular os pesos sinápticos das conexões entre os neurônios da rede. Em geral, procura-se carregar (codificar) tantos exemplos quanto for possível para dentro do modelo de rede treinado. Espera-se que a RNA assim projetada seja capaz de generalizar adequadamente. Generalizar adequadamente significa que o mapeamento entre os padrões de entrada e a saída obtida está correto tanto para os exemplos de treino (usados para geração da RNA), quanto para os exemplos de teste (usados para avaliação). Os exemplos de teste

são padrões retirados da mesma população dos exemplos de treino, mas que são inéditos para a RNA [19].

O processo de treinamento (aprendizagem) pode ser visto como um problema de ajuste de curva, sendo a rede um mapeamento não-linear de entradas-saídas. Esse ponto de vista permite considerar a generalização como uma boa interpolação não-linear dos dados de entrada [61] *apud* [19]. Quando o número de exemplos é infinitamente grande e os exemplos não são tendenciosos, os parâmetros internos da rede convergem para um dos mínimos locais da função de risco empírica (perda esperada). Quando o número de exemplos de treino é finito, a função de risco real (erro de generalização) é diferente da função de risco empírica [19].

Em resumo, quanto menor a quantidade de exemplos disponíveis, mais vulnerável o procedimento de treinamento está em relação às influências de outras características existentes no conjunto de exemplos de treino (*e.g.*, presença de ruído, tendências devido à proporção de exemplos de cada categoria). Além disso, o excesso de treinamento ou excesso de ajuste também prejudica o desempenho da rede. Ele ocorre quando os parâmetros internos de uma RNA se ajustam às particularidades dos exemplos de treino, devido a um prolongamento do procedimento de treinamento, e não são ótimos no sentido de minimizar o erro de generalização (*e.g.*, taxas de acerto sobre conjunto de teste) [62].

Determinar a parada do treinamento de uma RNA de modo a manter uma estrutura bem generalizada é complicado quando se considera somente o erro médio quadrático de treinamento [19]. Existem alguns métodos para evitar o excesso de treinamento e as influências de conjuntos de treinamentos reduzidos, como: métodos para seleção de modelos, parada antecipada, regularização, minimização do risco estrutural, e a validação cruzada. O procedimento de parada antecipada consiste em tentar interromper o treinamento da rede neural na época de maior generalização [19][62]. Essa interrupção apóia-se no procedimento de validação cruzada.

A validação cruzada consiste em dividir o conjunto de exemplos de treinamento em dois grupos: um conjunto de estimação (ou treino), e um conjunto de validação (ou teste). O conjunto de estimação tem por objetivo selecionar os parâmetros internos da rede neural, ou seja, treiná-la. Periodicamente esse treinamento é interrompido (geralmente ao fim de uma época) para avaliar o erro em relação ao conjunto de treino e de validação [19][62]. Heuristicamente, a curva do erro de treinamento tende a decrescer rapidamente, e então estabilizar-se num

ritmo de decrescimento menos acentuado. A curva de validação decresce mais lentamente que a de treinamento, até atingir um ponto mínimo, e então volta a crescer (Figura 11). As épocas próximas ao ponto de menor erro de validação apresentam as maiores taxas de generalização. Presume-se que a informação que a RNA absorve após esse ponto é essencialmente ruído presente nos dados [19][62].

O procedimento de parada antecipada consiste em parar o treinamento no ponto de menor erro sobre o conjunto de validação. A combinação desses dois procedimentos tem sido empregada no treinamento e avaliação de modelos de redes neurais e máquinas de vetor de suporte [16][26][60][63]. Apesar de delimitar um intervalo de épocas do treinamento que possuem as maiores capacidades de generalização, esses procedimentos não fornecem evidências para a escolha das épocas de maior desempenho.

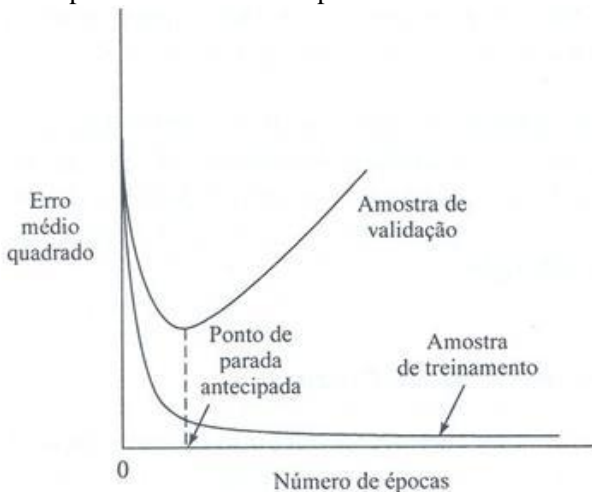


Figura 11. Curvas do erro médio quadrático de treinamento e de validação de uma RNA

Fonte: Adaptado de HAYKIN (2001, p.243) [19]

2.7. AVALIAÇÃO DO DESEMPENHO DE IDENTIFICAÇÃO

Os índices de desempenho quantificam a capacidade do classificador (rede neural artificial) em identificar os eventos de interesse. A avaliação e comparação do desempenho de RNAs na identificação de eventos têm usado os índices de sensibilidade,

especificidade, eficiência, e curvas ROC (*Receiver Operating Characteristic*) [64][65]. Esses índices advêm das áreas de Epidemiologia, Radiologia, e Bioestatística, onde são empregados na avaliação de exames clínicos de resposta binária (e.g., diagnóstico de patologias). Por exemplo: um observador deve registrar a ocorrência do comportamento de locomoção em segmentos de vídeos experimentais. Ele registra positivo para a ocorrência da locomoção, e negativo para a ocorrência de outros comportamentos. Têm-se as seguintes possibilidades de resposta [66]:

- Ao visualizar uma amostra de locomoção, o observador corretamente a registra como locomoção (Verdadeiro-positivo – VP);
- Ao visualizar uma amostra de outro comportamento, o observador erroneamente o marca como locomoção (Falso-positivo - FP);
- Ao visualizar uma amostra de outro comportamento, o observador a rotula como negativo (Verdadeiro-negativo - VN); e
- Ao visualizar uma amostra de locomoção, o observador erroneamente a marca como outros (Falso-negativo - FN).

O índice de sensibilidade (ou taxa de verdadeiro-positivos) representa a capacidade do procedimento de identificar corretamente as amostras do evento-alvo (e.g., locomoção). Um método que possui um baixo número de FN é dito de alta sensibilidade [16][60][66][67]. O cálculo da Sensibilidade está descrito na Equação 1.

$$sensibilidade = \frac{VP}{VP + FN} \quad (1)$$

O índice de especificidade retrata a capacidade de identificar corretamente as amostras de eventos não-locomotores (VN). Baixas taxas de FP indicam métodos de alta especificidade. O cálculo da especificidade está descrito na Equação 2 [16][60][66][67]:

$$especificidade = \frac{VN}{VN + FP} \quad (2)$$

O relacionamento entre os níveis de sensibilidade e especificidade no plano cartesiano é conhecido como Curva ROC (vide Figura 12) [66][60]. O estudo das curvas ROC teve início em 1940 nos campos de psicologia e sistemas eletrônicos de comunicação. Atualmente essas curvas têm sido bastante empregadas para avaliar o desempenho de diagnóstico de sistemas especialistas e de redes neurais artificiais. Em redes neurais artificiais, os pontos que formam a curva ROC são obtidos pela medição da sensibilidade e da especificidade do neurônio de saída da rede usando diferentes pontos de corte da função de ativação (e.g., 0,1; 0,2; 0,3;...; 0,8; 0,9)[16][65][66][68].

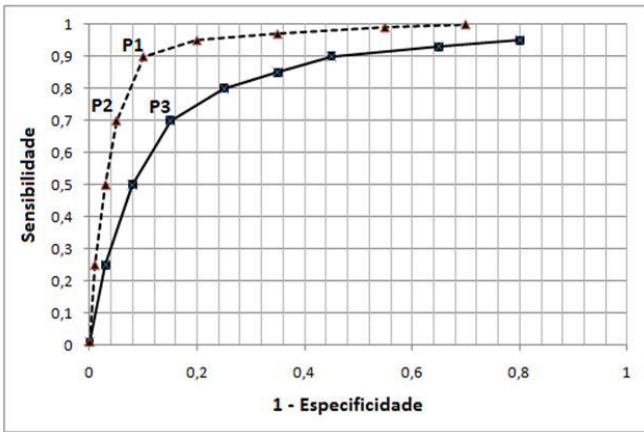


Figura 12. Comparação de desempenho de dois classificadores através de suas curvas ROC. Os pontos P1, P2 e P3 representam o desempenho de diagnóstico (em pares de sensibilidade e especificidade) de diferentes pontos de corte da função de ativação do neurônio de saída de classificadores de resposta binária [68][66].

Quando as curvas ROC de um conjunto de classificadores são muito próximas, e tanto o formato da curva quanto os índices de sensibilidade e especificidade são semelhantes, a simples observação das curvas não permite a identificação do elemento de maior desempenho. Nesses casos calcula-se a área sob a curva ROC (AUC). Uma AUC de 1 (100%) sugere um teste que obteve o desempenho máximo, e sua sensibilidade é 100% e a taxa de falso-positivos (1 – especificidade) é zero, enquanto um teste de AUC zero é totalmente ineficiente.

A eficiência também expressa o relacionamento entre os índices de sensibilidade e especificidade em apenas uma medida [64][69].

Descreve-se como a razão entre a soma dos acertos nos índices VP e VN pela soma de todos os eventos analisados (VP, VN, FP, FN), vide Equação 3.

$$eficiência = \frac{VP + VN}{VP + VN + FP + FN} \quad (3)$$

Os índices descritos são recomendados quando existe um padrão ouro de resposta [64]. Em situações em que não existe esse padrão, pode-se tomar como referência as respostas de um especialista da área. Contudo, essa decisão assimilará os vieses desse registro. Existe índices voltados a avaliação da concordância entre especialistas (ou observadores). Esses índices se aplicam principalmente aos casos nos quais a própria natureza dos eventos analisados é ambígua (*e.g.*, registro comportamental). Nesses casos a decisão de que um evento pertence (ou não) a uma determinada categoria é determinada pelo consenso de diagnóstico de dois ou mais especialistas. Dois índices comuns na avaliação da concordância entre observadores nas áreas da Etologia e Ciências comportamentais são o índice de concordância (IC), e o coeficiente Kappa [2][5].

O IC estima a concordância bruta (ou pura) entre dois observadores. É apropriado para medidas nominais ou classificatórias. Por exemplo: ao fim de uma sessão de observação, dois observadores comparam o número de concordâncias (C) e discordâncias (D) entre seus registros. O IC é a razão entre o número de concordâncias sobre o número total de ocorrências (C+D), conforme Equação 4 [2].

$$IC = \frac{C}{C + D} \quad (4)$$

O coeficiente Kappa é uma especialização do IC. Ele subtrai de IC uma estimativa do número de concordância que ocorram por acaso (índice *ac*). *ac* é a probabilidade de que ambos observadores registrem um evento comportamental (*e.g.*, de locomoção) para a mesma amostra, dado que seus registros são distribuídos aleatoriamente. O cálculo do índice Kappa é descrito pela Equação 5.

$$k = \frac{IC - ac}{1 - ac} \quad (5)$$

Onde:

$$ac = \sum_1^{N_c} \left[\left(\frac{ev_{j,obs1}}{N} \right) * \left(\frac{ev_{j,obs2}}{N} \right) \right] \quad (6)$$

$ev_{j,obs1}$: quantidade de eventos j registrados pelo observador

N : número de amostras

N_c : número de comportamentos

Por exemplo, comparam-se dois observadores (1 e 2) que tem por objetivo registrar a frequência de um padrão de comportamento (*e.g.*, imobilidade) por 15 minutos, com amostragem temporal instantânea em intervalos de 30 segundos. Os observadores devem marcar em quais das trinta (30) amostras consideram que o comportamento de imobilidade ocorreu. Verificou-se que o observador 1 registrou o comportamento de imobilidade em 16 amostras, e o observador 2 em 13. Assume-se como verdade que ambos concordaram no diagnóstico em 25 das 30 amostras visualizadas (Locomoções e Não-locomoções). O índice de concordância desse caso foi estimado em 83% (25/30) [2]. A estimativa da proporção de concordâncias por acaso é demonstrada abaixo, usando a Equação 6.

$$ac = (16/30 * 13/30) + (14/30 * 17/30). \\ ac = 0,49.$$

Substituí-se o valor obtido de IC e ac na (5, e obtém-se o valor de Kappa.

$$k = (0,83 - 0,49) / (1 - 0,49) = 0,67.$$

De acordo com o valor obtido do coeficiente Kappa, a concordância real entre os observadores foi de 67%. Ou seja, 16% das concordâncias apresentadas pelo IC foram por acaso, e não por um real conhecimento do observador [2]. O índice Kappa já foi usado para auxiliar a determinação do ponto de corte de funções de ativação de RNA MLP e na comparação de desempenho entre essas redes e observadores humanos [63][71][72].

Os índices de avaliação do desempenho de diagnóstico e da concordância entre observadores são geralmente associados a procedimentos de avaliação da generalização, como critérios na comparação e seleção entre modelos de redes neurais treinados. No entanto, eles não avaliam se os parâmetros de entrada usados são relevantes para a descrição dos eventos de interesse. Como esta tese

estudou descritores de eventos comportamentais, é de particular interesse avaliar quais desses descritores são relevantes para a detecção dos eventos de interesse.

2.8. PRÉ-SELEÇÃO DOS PARÂMETROS DE ENTRADA

Métodos de seleção (ou pré-seleção) de parâmetros de entrada auxiliam a selecionar o subconjunto de parâmetros mínimo (ou ótimo) que maximize a capacidade de detecção de um método de classificação [73][74]. Eles são uma alternativa a busca exaustiva, que apesar de garantir o melhor conjunto de parâmetros, torna-se impraticável a medida que o número de parâmetros aumenta.

Os procedimentos de seleção geralmente englobam um método gerador de um subconjunto de atributos de teste, uma função de avaliação, um critério de parada, e um procedimento de avaliação (ou validação). O método gerador do subconjunto pode iniciar com um conjunto vazio, completo, ou com um subconjunto aleatório. A função de avaliação mede a qualidade ou desempenho de um subconjunto, e permite ao procedimento de seleção comparar e selecionar os subconjuntos gerados. O critério de parada evita que o procedimento de busca seja conduzido de forma exaustiva e indefinida. A etapa de validação não faz parte do procedimento de seleção em si, mas é necessária para averiguar a aplicabilidade do método para um determinado domínio [75].

Os métodos de seleção são classificados segundo as funções que empregam para a avaliação da relevância dos parâmetros (medidas de distância, da informação, dependência, consistência, e erro de classificação). Os métodos que se baseiam no erro de classificação (conhecidos como *wrappers*), usam do próprio algoritmo de classificação automática que fará uso dos parâmetros para avaliar os subconjuntos de parâmetros candidatos. Apesar de computacionalmente custosos, os *wrappers* apresentam uma alta taxa de acerto [73][75]. Em geral, eles iterativamente incrementam ou decrementam o número de parâmetros dos subconjuntos candidatos até encontrar uma solução ótima [16][23].

Nos procedimentos de pré-seleção conhecidos como “filtros”, os parâmetros de entrada candidatos são avaliados e ordenados de acordo com uma função de avaliação, que determina sua relevância para tarefa. Esses métodos produzem resultados mais rápidos que os métodos *wrappers*, porém ignoram interações entre os parâmetros, e não

identificam parâmetros redundantes [73]. Exemplos desses métodos são os testes de chi-quadrado [76], e de Wilcoxon-Mann-Whitney [77].

Soluções híbridas que iniciam a seleção dos parâmetros com o uso de procedimentos do tipo filtro, e a partir de um conjunto reduzido de parâmetros candidatos, avaliam a relevância dos elementos usando métodos de classificação automática. A escolha do método a empregar (*wrapper*, filtro, ou híbrido) varia de acordo com o domínio do problema e as restrições dos sistemas desenvolvidos.

Em resumo, um classificador treinado a partir de um conjunto de atributos que são irrelevantes para representação do fenômeno de interesse será ineficaz em produzir bons resultados, independente do índice que for usado para avaliar seu desempenho. Portanto, além de empregar índices que retratem o desempenho de um modelo de classificador e sua capacidade de generalização, é factível empregar um procedimento de pré-seleção para avaliar os parâmetros usados (2º objetivo específico). A eliminação de parâmetros desnecessários aumenta a habilidade de generalização de modelos de RNA [17].

3. MATERIAIS E MÉTODOS

Neste capítulo são descritos os procedimentos empregados para obtenção e estudo do sinal comportamental. A execução do procedimento experimental foi conduzida pelo próprio autor desta tese, após participar de um curso de bioterismo e ética no uso de animais. Os experimentos foram conduzidos com colaboração de alunos da pós-graduação em Neurociências no laboratório de Bioengenharia do Instituto de Engenharia Biomédica da Universidade Federal de Santa Catarina (IEB-UFSC). Todos os procedimentos adotados foram previamente submetidos e aprovados pela Comissão de Ética no Uso de Animais (CEUA) da UFSC. O conjunto de comportamentos estudados está descrito na seção 3.1.5. O registro comportamental foi realizado após a execução do experimento, usando as gravações em vídeos digitais dos experimentos. Toda etapa de registro comportamental e análise de atividade usou a ferramenta desenvolvida durante esta tese, o ETHOWATCHER® (descrita na seção 3.2). Os descritores morfológicos e cinemáticos foram extraídos dos segmentos de vídeo de cada categoria comportamental (conforme registro do observador dos experimentos, seção 3.3). Com intuito de pré-selecionar os atributos estatísticos relevantes para a diferenciação e reconhecimento dos eventos comportamentais, empregou-se a prova estatística não paramétrica de *Kruskal-Wallis*. A validação dos descritores e do método de pré-seleção se baseou na comparação do desempenho obtido pelo uso de diferentes conjuntos de descritores e atributos estatísticos como padrões de entrada em modelos de RNA MLP (seção 3.6).

3.1. PROTOCOLO EXPERIMENTAL

3.1.1. Ambiente Experimental

O ambiente de experimentos foi uma arena do tipo campo aberto, em plástico, de formato circular, com altura de 45 cm, e raio de 50,5 cm, pintada internamente com tinta acrílica preta. O ambiente estava iluminado diretamente por uma lâmpada fluorescente de 16 W (313 lux no centro da arena). A aquisição dos vídeos foi realizada por uma filmadora SONY DCR-SR45 (DCR-SR45 Sony, USA - configurada para adquirir vídeos a 30 quadros por segundo) posicionada

perpendicularmente ao centro da arena. Posteriormente os vídeos foram convertidos do formato de aquisição (MPEG2) para o formato *Audio Video Interleaved* (AVI, formato nativo do Windows Media Player, usado pela ferramenta ETHOWATCHER®), usando o software VirtualDub (version 1.78, <http://www.virtualdub.org>). A resolução dos vídeos foi reduzida para 320x240 pixels (suficiente para uso de webcams de baixo custo), e do sistema de cores RGB para escala de níveis de cinza de oito bits. Por fim os vídeos foram comprimidos usando o codec XVID MPEG4.

3.1.2. Animais

Foram utilizados ratos *Rattus norvegicus* (Wistar), fêmeas, com peso corporal variando entre 250-300g. Os animais foram mantidos em caixas de polipropileno (49 x 34 x 16 cm) forradas com maravalha, sendo submetidos a sete dias de adaptação ao biotério do laboratório antes dos testes experimentais. Disponham de água e alimento granulado (CR-1, Nuvilab) ad libitum, em uma sala com temperatura controlada entre 22 °C a 24 °C, e um ciclo claro-escuro de 12 horas (luz das 7-19 h). O protocolo experimental usado foi aprovado no dia 16/07/2008 pela Comissão de Ética no Uso de Animais (CEUA - protocolo PP00254 - Ofício nº 128/CEUA/PRPe/2008).

3.1.3. Soluções e doses administradas

As doses escolhidas se baseiam em estudos anteriores, que as indicaram como provocadoras de aumento na atividade locomotora em roedores, alterações as quais foram medidas pela interrupção de sensores luminosos verticais e horizontais [78][79][80], ou o registro do sinal produzido pela atividade motora sob sensores de movimento [81]. As seguintes drogas e doses foram utilizadas:

- Cafeína (1,3,7-trimethylxanthine, RBI, Natick, MA, U.S.A.), dissolvida em solução aquosa de benzoate de sódio 10 mg/ml nas doses de 2,0 (dose 1) e 6,0 mg/Kg (dose 2), aplicadas em um volume de 1 ml (ip).
- Salina: Solução de NaCl (0.9% estéril, livre de pirogênicos, pH 7,4; (1 ml, ip).

A preparação e a aplicação dessas doses foram acompanhadas por mestrandos do curso de pós-graduação em Neurociência (CCB-UFSC). A alteração da expressão comportamental foi necessária para avaliar a capacidade de generalização das redes neurais treinadas, quando em face de amostras de comportamentos provenientes de animais tratados com drogas que alteram a expressão comportamental.

3.1.4. Administração das Drogas

Nestes experimentos seis (6) ratos experimentalmente ingênuos a tratamentos receberam injeções sistêmicas de cafeína ou de salina (controle), 10 minutos antes de serem colocados na arena de campo aberto. A atividade desses animais no campo aberto foi registrada por pelo menos 10 minutos após sua introdução no ambiente. Os animais foram distribuídos em três grupos, e a cada rodada recebiam uma das doses de cafeína ou de salina (controle). Todos os animais receberam todos os tratamentos (de acordo com um esquema amostral de quadrado latino, vide Quadro 2), com pelo menos seis dias de intervalo entre cada tratamento.

Animal	Rodada 01	Rodada 02	Rodada 03
Rato 01	Salina	Cafeína 6 mg/kg	Cafeína 2 mg/kg
Rato 02	Salina	Cafeína 6 mg/kg	Cafeína 2 mg/kg
Rato 03	Cafeína 2 mg/kg	Salina	Cafeína 6 mg/kg
Rato 04	Cafeína 2 mg/kg	Salina	Cafeína 6 mg/kg
Rato 05	Cafeína 6 mg/kg	Cafeína 2 mg/kg	Salina
Rato 06	Cafeína 6 mg/kg	Cafeína 2 mg/kg	Salina

Quadro 2. Quadrado latino da sequencia de administração das drogas em cada animal

Ao fim dos experimentos os animais foram anestesiados profundamente com uma mistura de Ketamina e Xylazina (0,15 e 0,05 ml/100 g de peso corporal, respectivamente, ip), e sacrificados por perfusão transcordial com salina seguida por formol a 10%.

3.1.5. Catálogo de Comportamentos

Visto que o estudo da detecção automática foi restrito ao modelo animal de campo aberto, selecionaram-se alguns comportamentos comuns às pesquisas empregando esse modelo animal [34]. Seguem as descrições de cada categoria comportamental, complementadas por imagens que caracterizem sua expressão.

- **Locomoção:** trata-se do deslocamento do animal com a respectiva alteração de posição de suas quatro patas [82]. Ou, o movimento do animal de um lugar para outro, que procura refletir uma distancia mínima percorrida nesse deslocamento [83]. Nas imagens da Figura 13 o animal modifica tanto sua posição quanto sua direção durante a expressão deste comportamento;

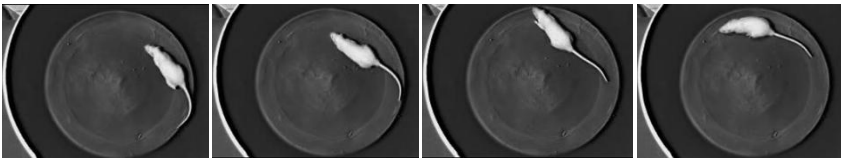


Figura 13. Locomoção a cada 10 quadros

- **Imobilidade:** trata-se do animal praticamente imóvel. Admitem-se pequenos movimentos da cabeça e tremores. Nos quadros ilustrados na Figura 14, praticamente não há nenhuma mudança na posição ou forma do animal;

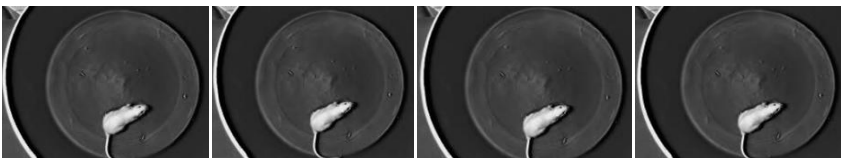


Figura 14. Imobilidade a cada 10 quadros

- **Auto-limpeza, limpeza corporal ou manutenção:** quando o animal apresenta comportamentos, como: lavar a cabeça, limpar as unhas, coçar o corpo. É conhecido na literatura pelo termo em inglês *grooming*. Nos quadros da Figura 15, verifica-se poucas alterações na forma da área do animal. Essas alterações restringem-se a região

da cabeça, local no qual o animal está realizando o ato de limpar. Um leve encurtamento do corpo do animal pode ser notado quando comparado com o comportamento de imobilidade. O *grooming* remanesce como um dos comportamentos de mais difícil detecção tanto para sistemas de detecção baseados em processamento de vídeos, como de sensores luminosos e de vibração [14]. Resultados preliminares de sistemas que fazem uso de descritores de Fourier e redes neurais de aprendizado não supervisionado demonstraram que esses modelos são capazes de identificar essa categoria de comportamento [21]

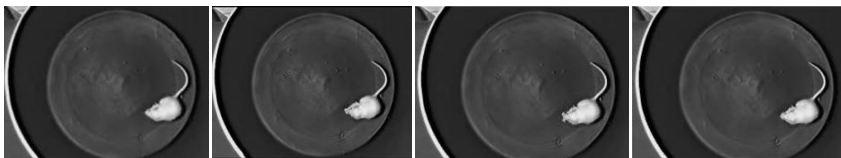


Figura 15. Limpeza corporal a cada 10 quadros

- Exploração Vertical: quando o animal estiver explorando o ambiente apoiado somente nas patas traseiras, com as patas dianteiras livres, e não apoiadas no fundo da caixa (ereto), conforme Figura 16. É conhecido na literatura pelo termo em inglês *rearing*.

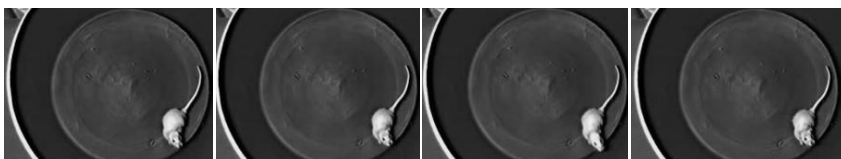


Figura 16. Exploração Vertical a cada 10 quadros

Conforme exemplos (Figuras 13-16), a representação dos comportamentos não se resume a apenas uma imagem. A visualização individual de imagens de comportamentos citados, como a locomoção (Figura 13) e a imobilidade (Figura 14) não permite as diferenciar, pois são praticamente idênticas. A limpeza corporal (Figura 15) difere desses comportamentos pelo animal estar mais retraído, evidência também comum em amostras de imobilidade. É a variação temporal do contorno da área do animal ou de seu deslocamento cinemático que elimina as

ambiguidades existentes entre os comportamentos, e caracteriza uma categoria.

As definições do catálogo de comportamentos descrito foram usadas para registrar (de forma manual) os eventos comportamentais dos experimentos realizados (usando a ferramenta ETHOWATCHER®). A Tabela 1 ilustra a frequência de cada categoria comportamental por grupo de tratamento.

Tabela 1. Frequência dos comportamentos por grupo de tratamento

Comportamentos	Veículo	Cafeína 2 mg/kg	Cafeína 6 mg/kg	Total
Locomoção	40	31	48	119
Imobilidade	110	36	114	260
Exploração Vertical	60	57	100	217
Limpeza corporal	20	13	6	39

O protocolo experimental restringiu o registro de comportamentos a eventos com duração igual ou superior a 1 segundo (equivalente a 30 quadros para a filmadora em uso). Essa restrição deve-se à dificuldade em se caracterizar com precisão eventos de duração inferior ao tempo mínimo definido.

O estudo dos descritores desenvolvidos iniciou com a extração dos descritores dos segmentos de cada vídeo experimental que haviam sido marcados como eventos comportamentais de interesse. Em resumo, executava-se um procedimento de análise automática da atividade animal para cada vídeo, como realizado por estudos comportamentais e farmacológicos (usando ferramentas de rastreamento animal), porém extraíram-se as variáveis cinemáticas (*e.g.*, distância percorrida, velocidade média, e número de pixels alterados) e morfológicas (*e.g.*, área e comprimento do animal) apenas dos segmentos de tempo marcados como comportamentos. O agrupamento dos descritores comportamentais por categoria comportamental permitiu avaliar quais os descritores que as discerniam umas das outras. Optou-se por desenvolver uma ferramenta própria (o ETHOWATCHER®) para realizar o procedimento de extração descrito, já que nenhuma ferramenta encontrada era capaz de fazê-lo.

3.2. ETHOWATCHER®

A necessidade de uma ferramenta que permita desde o registro até a análise da atividade dos eventos comportamentais foi apenas uma das razões para o desenvolvimento do ETHOWATCHER®. A análise

de atividade empregada pelas ferramentas existentes se restringe ao rastreamento da atividade do animal ao longo do experimento (*video-tracking*), e o processamento de dados relativos a essa atividade, como: velocidade média, distância percorrida, e tempo de permanência em determinados quadrantes da arena comportamental. A descrição de padrões comportamentais exige o teste de diferentes conjuntos de variáveis (leia-se índices) relativas à análise de atividade animal e a morfologia de sua forma (geralmente incomuns a análise padrão de ferramentas comerciais). Além disso, os dados de atividade geralmente são expressos em termos de soma, média, desvio padrão e erro padrão. O acesso aos valores individuais de cada quadro dos segmentos de vídeo comportamental é essencial, pois é a partir desses dados que podem ser testados diferentes métodos de descrição do comportamento. O uso das ferramentas existentes inviabilizaria o acesso a essas informações. Esses são alguns dos argumentos que induziram o desenvolvimento da ferramenta ETHOWATCHER®. Além disso, o laboratório de Bioengenharia do IEB-UFSC já possuía ferramentas individuais que contemplavam as etapas de análise da atividade [42] e registro comportamental [29], que embora não validadas em situação experimental real, contribuíram significativamente para o desenvolvimento do ETHOWATCHER®. A seguir serão descritos os procedimentos de processamento de imagens empregados pela ferramenta, previamente a extração dos descritores.

O procedimento de análise de atividade (cerne do ETHOWATCHER®) baseia-se no modelo construído por BOSE [42], adaptado de Intel [84]. Neste modelo o processo de análise de atividade é dividido em duas etapas. A primeira etapa é supervisionada pelo usuário, e trata da calibração do algoritmo de segmentação. O usuário seleciona uma imagem da cena experimental vazia (que deve estar presente no arquivo de vídeo), e a partir de uma cena qualquer contendo o animal, testa diferentes valores de limiar, com objetivo de minimizar a presença de artefatos (ruído) na imagem resultante da subtração da imagem da cena vazia com a cena contendo o animal. A segunda etapa desse modelo executa o procedimento de análise da atividade. A análise da atividade animal é composta da etapa de segmentação do animal do plano de fundo com base nos parâmetros da etapa de calibração, e da extração das variáveis cinemáticas e morfológicas do objeto segmentado (*e.g.*, rato, camundongo).

O procedimento de segmentação adotado foi a subtração de plano de fundo por imagem estática [18][42][84]. Esse procedimento subtrai uma imagem de referencia (cena experimental sem a presença do

animal) da imagem contendo o objeto a segmentar. Seu uso assume que tanto a posição da câmera quanto a iluminação de ambiente sejam relativamente constantes. Desta forma, espera-se que os únicos pixels diferentes na cena em análise sejam o objeto a segmentar, logo a imagem resultante da subtração conterá apenas o novo objeto de interesse. As pequenas variações de intensidade luminosa (que resultam em pixels isolados não pertencentes ao objeto) são retiradas através da aplicação da técnica de Limiarização na imagem resultante da subtração de imagens. Intensidades de cor inferiores a esse limiar são removidas da imagem resultante.

Variações na intensidade de cor da cena ao longo do experimento (*e.g.*, devido a falhas de alimentação da iluminação ambiente) produzem ruídos que o método em uso é incapaz de remover, e são rotuladas como pixels do objeto a segmentar. Para contornar essa vulnerabilidade foi introduzido um fator de atualização da cena vazia. Executa-se uma amostragem das diferenças de intensidade entre 50 pixels da imagem em análise em relação à cena vazia. A média dessas diferenças constitui o fator atualização. Com esse incremento, para que um determinado pixel seja rotulado como pertencente ao objeto de interesse, sua intensidade deve estar acima da soma do limiar definido pelo usuário na etapa de calibração com o fator de atualização calculado. Esse procedimento retirou grande parte dos ruídos de iluminação presentes nos vídeos experimentais.

Por último foi adicionado ao procedimento de segmentação um filtro de intensidade de cor [48]. Após a aplicação do limiar de segmentação e do fator de atualização, a matriz resultante é testada pixel a pixel quanto à intensidade de cor de esses pixels pertencerem a um intervalo de intensidades de cor definido durante a calibração. Esse procedimento eliminou a influência de ruídos na imagem resultantes da presença de fezes e urina do animal na arena. Os pixels existentes nessa matriz são admitidos como parte da área do animal. Pixels do ambiente são marcados durante esses procedimentos com uma cor definida como neutra (cor escolhida pelo usuário como a cor do ambiente experimental, durante a calibração do sistema). A cor neutra será ignorada pelos algoritmos subsequentes de análise de atividade e extração de descritores de imagens.

O módulo de análise de atividade permite o rastreamento de um animal por vez, e trabalha com vídeos na resolução de 320x240 pixels, no formato AVI, em escala de cinza. Essas configurações tornam-no compatível com diferentes categorias de dispositivos de aquisição de imagens (de câmeras web a filmadoras). É compatível com arenas de

campo aberto, labirinto aquático de Morris e labirinto elevado em cruz (vide exemplos desses ambientes na seção 2.1). Ao fim do processamento da atividade é produzida uma imagem da cena vazia com a ilustração do caminho percorrido pelo animal na arena durante o experimento (vide exemplos da Figura 6 na seção 2.2), e um relatório descrevendo quadro a quadro, as variáveis cinemáticas e morfológicas extraídas (*e.g.*, número de pixels do animal, posição do centróide do animal, comprimento médio do animal).

O outro módulo do ETHOWATCHER® auxilia o registro dos comportamentos (*e.g.*, de animais, de pessoas) por um observador humano. Nesse caso um observador treinado cria previamente um catálogo de comportamentos no software, e associa uma tecla de atalho a cada categoria comportamental a registrar. A partir desse catálogo, o pesquisador registra os comportamentos pela observação do animal em tempo real, ou de arquivos de vídeos. Ao fim do procedimento de registro a ferramenta gera um relatório com todos os comportamentos registrados na sequência temporal original, e um sumário contendo a frequência, a duração total, e a latência de cada categoria do catálogo.

Conforme discutido, um dos principais motivos de desenvolver esse sistema foi a liberdade de acesso aos dados brutos da análise de atividade, e implantação dos descritores desejados nesse processo. Essa liberdade permitiu customizar a análise da atividade animal (e extração dos descritores) para extrair os dados apenas dos segmentos de vídeo marcados como um comportamento durante a Etografia (usando o módulo próprio do ETHOWATCHER®). Em resumo, uma segmentação comportamental da análise de atividade do animal. Na prática o sistema produz um arquivo de texto para cada comportamento registrado no vídeo experimental, sendo que cada arquivo contém todos os descritores medidos em cada quadro do respectivo segmento de tempo. Os arquivos de texto são formatados no padrão CSV (*Comma-separated values*), compatível com a maioria dos softwares de planilha de cálculo disponíveis (*e.g.*, STATISTICA, STATSOFT, versão 8.0).

O ETHOWATCHER® foi registrado no Instituto Nacional Propriedade Intelectual (INPI) sobre o protocolo 09285-3, na revista eletrônica Nº 1982, com regime de guarda de sigilo até 13/11/2018. O software está sendo distribuído de forma gratuita na Internet no endereço: www.ethowatcher.ufsc.br. As funções básicas da ferramenta estão descritas no APÊNDICE I – ETHOWATCHER. Descrições resumidas deste software foram objeto de apresentações de pôsteres em congressos nacionais e internacionais (I NEUROLATAM, SBNEC

2010), e sua descrição pormenorizada foi enviada para publicação no periódico indexado *Computer in Biology and Medicine* (em avaliação).

Na próxima seção serão descritos os índices cinemáticos e morfológicos implementados, automaticamente extraídos pelo ETHOWATCHER®. Esses descritores foram validados em situações experimentais reais, nas quais as respostas dos animais frente ao tratamento eram conhecidas, conforme procedimento descrito no APÊNDICE B.

3.3. DESCRITORES COMPORTAMENTAIS

O uso de descritores morfológicos é visto na literatura como um meio de capturar as características essenciais das categorias comportamentais, e usá-las como parâmetros de entrada em máquinas de aprendizado para classificação automática. Em geral, esses descritores avaliam as características do contorno ou da forma da área do animal (e.g., a área do animal, o centro de massa, o formato do animal, a distância de seu centróide até seu focinho ou até sua cauda) [17][21][43][44][85]. A descrição comportamental através da componente cinemática do comportamento é rara [16], e tem sido discutida como uma alternativa mais adequada para a descrição de algumas categorias comportamentais [15].

Por exemplo, algumas imagens intermediárias de uma locomoção, quando vistas isoladamente assemelham-se a imagens de animais imóveis. É a componente cinemática do movimento das partes (ou de todo) corpo no tempo que retira essas ambiguidades.

Os descritores morfológicos extraídos foram a área e o comprimento do animal, e os descritores cinemáticos foram a distância percorrida, o número de pixels alterados, e o ângulo de direção do animal em relação ao eixo vertical da imagem (vide Quadro 3). O cálculo dos descritores cinemáticos é uma relação entre o quadro atual e o anterior, enquanto os morfológicos referem-se apenas ao quadro em análise. Todos os descritores são calculados sobre os dados das matrizes obtidas após o procedimento de segmentação do objeto de interesse do cenário.

Todos os descritores foram extraídos através do ETHOWATCHER®, e foram calculados sobre a matriz resultante do procedimento de segmentação (obtida pela subtração de plano de fundo, e subsequentes ajustes descritos na seção 3.2). Essa matriz será citada nos parágrafos e seções a seguir como matriz resultante da segmentação.

Um comparativo entre a imagem original e a imagem resultante do procedimento é ilustrado na Figura 17 (A: original, B: resultado).

Descritores Comportamentais	Categoria
Área	Morfológico
Comprimento do Animal	Morfológico
Número de Pixels alterados	Cinemático/Morfológico
Distância Percorrida	Cinemático
Ângulo de inclinação da reta do comprimento do animal	Cinemático
Varição angular	Cinemático

Quadro 3. Descritores de imagem analisados

O primeiro parâmetro calculado é a posição estimada do corpo do animal no plano cartesiano (círculo cinza na Figura 17B). Para obtê-la efetuou-se a média das coordenadas cartesianas dos n pixels da imagem que foram rotulados como pertencentes ao corpo do animal (vide Figura 17B, e Equação 7) (adaptado de INTEL [84] por BOSE [42]). Entendem-se como pixels pertencentes ao corpo do animal, todos os pixels da matriz resultante da segmentação cuja intensidade de cor seja diferente da cor definida como neutra. A distância euclidiana entre a posição estimada do centro da área do animal no quadro atual e a posição estimada no quadro anterior constitui o descritor da distância percorrida pelo animal [18][42].

$$\begin{cases} u_x = \frac{\sum x}{TP} \\ u_y = \frac{\sum y}{TP} \end{cases} \quad (7)$$

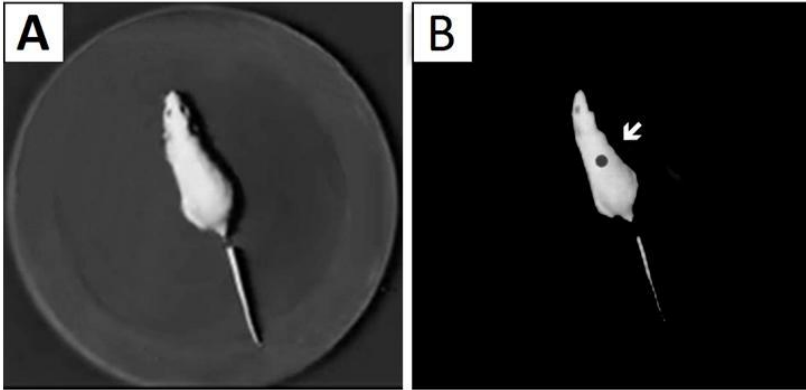


Figura 17. Imagem de um rato no ambiente de campo aberto (A), seguida da imagem resultante do procedimento de subtração, com ilustração do ponto estimado como centro da área do animal (B).

O descritor comprimento do animal obtém a maior reta que atravessa longitudinalmente a área do animal, e que cruza a posição estimada do centro da área de seu corpo no quadro em análise. Para tanto, emprega-se um procedimento de busca interativo que traça retas partindo do centro da área do animal, inclinadas em relação ao eixo das ordenadas no intervalo de ângulos de -1 a -89° e de 0 a 89° graus (para um incremento de 2 , têm um total de 179 retas, vide Figura 18B). Cada reta é incrementada em ambos os sentidos, até que ambas as extremidade ultrapasse a borda da área do objeto (vide Figura 18A). Considera-se ultrapassada a borda da área do objeto quando o novo pixel adicionado a extremidade da reta possui a intensidade de cor neutra. A média de comprimento das 10 maiores retas encontradas é admitida como o comprimento do animal. A média dos ângulos de inclinação dessas retas é admitida como a orientação do animal. A diferença entre a orientação do animal no quadro atual, e a orientação no quadro anterior é usada como o descritor da variação angular.

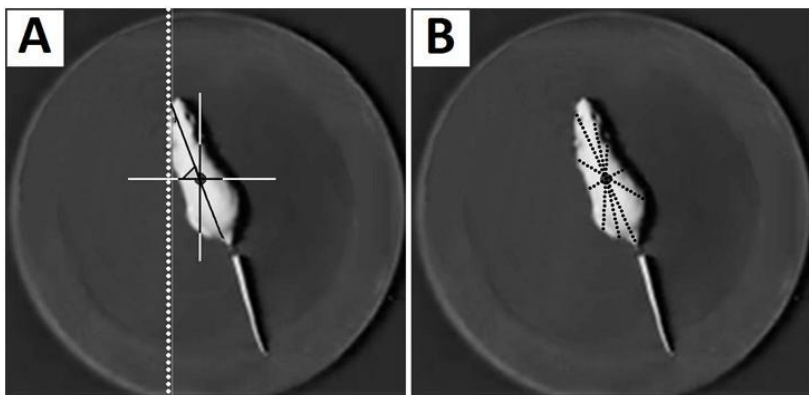


Figura 18. Método de Estimação do Comprimento do Animal. A Figura A ilustra o cálculo de uma reta i dado o ângulo α . A Figura B ilustra o conjunto de retas produzido pelo método para um intervalo de ângulos de -89 a 90 .

O descritor de número de pixels do animal quantifica o número de pixels rotulados como área do animal (todos os pixels da matriz resultante da segmentação cuja intensidade de cor seja diferente da cor neutra, Figura 19A). Alterações na área do objeto são medidas pelo descritor de número de pixels alterados (Figura 19B). Esse segundo descritor avalia diferenças na intensidade de cor (níveis de cinza) de cada pixel, na comparação da matriz resultante da segmentação do quadro atual em relação à matriz do quadro anterior. Trabalhos anteriores convertiam em valores binários as intensidades de cor da área do animal, e comparavam se haviam ocorrido alterações espaciais dos pixels do animal entre a imagem em análise e a imagem anterior [16] [18]. O descritor aqui implementado analisa tanto a variação espacial dos pixels que pertencem à área do animal, quanto as modificações locais da intensidade de cor dos pixels. Essas alterações de cor podem guardar informação comportamental relevante, como dados para identificação de um movimento fisiológico (respiração), ou para a diferenciação de comportamentos semelhantes (imobilidade e limpeza corporal).

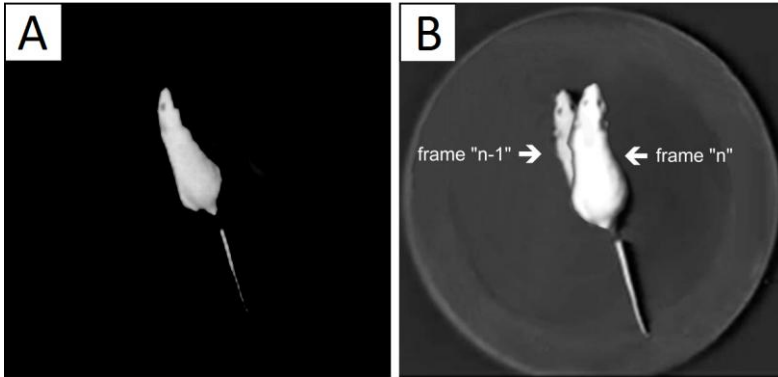


Figura 19. A área do animal é estimada como a quantidade de pixels com intensidades de cor diferente do plano de fundo na imagem resultante da segmentação (A). O número de pixels alterados é a quantidade de pixels que mudaram de intensidade de cor na comparação entre as matrizes pós procedimento de segmentação referentes a imagem (frame) atual (n) e da anterior (n-1). A figura B ilustra essa comparação sobrepondo a informação pré-segmentação de dois quadros consecutivos.

Um exemplo de relatório do módulo de análise da atividade do *software* ETHOWATCHER® (contendo os descritores apresentados) está presente no APÊNDICE A.

3.4. ATRIBUTOS DE ESTATÍSTICA DESCRITIVA

Um observador humano ao visualizar um vídeo experimental, naturalmente armazena e relaciona os padrões comportamentais das diversas imagens que está visualizando, e quando convencido, registra um segmento de vídeo (e não apenas uma imagem) como uma ocorrência da(s) categoria(s) em vista. É comum para esses especialistas avançar e retroceder um vídeo experimental para confirmar a ocorrência de uma categoria, e decidir seu início e fim aproximados. Essa aproximação é um dos maiores vieses do registro comportamental, pois dificulta a obtenção de um resultado padrão entre diferentes laboratórios [15][86].

Os índices morfológicos e cinemáticos procuram quantificar os padrões comportamentais usados pelo observador para registrar os eventos. No entanto, esses descritores extraem a informação referente a uma imagem, ou no máximo de um par delas. Como a informação de uma única (ou de um par de) imagem é insuficiente para identificar a

ocorrência da maioria das categorias de forma precisa, faz-se necessário usar de métodos complementares que representem a informação contida no conjunto de imagens que constitui um evento comportamental.

Entende-se que é a variação (ou invariância) dos índices durante as imagens contidas em um comportamento que contribui para a identificação (ou diferenciação) precisa entre categorias de comportamentos. Logo, é importante que essa variação seja descrita e incluída no procedimento de descrição e reconhecimento comportamental. Essa associação entre as imagens que compõem um evento comportamental é mais evidente em comportamentos estereotipados (*e.g.*, exploração vertical). Nesses comportamentos uma sequência de posturas em uma determinada ordem configura a ocorrência do evento comportamental.

A descrição desse relacionamento pode ser feita usando descritores de Fourier [21], heurísticas e regras de decisão sobre a sequência de posturas exibidas pelo animal [21][26][87], ou agrupavam os descritores referentes a um intervalo de tempo considerado suficiente, e os usava diretamente como padrões de entrada de um método de classificação automática (*e.g.*, aumentando o número de neurônios na camada de entrada desses modelos) [17].

Optou-se por usar atributos de estatística descritiva para representar a variação dos descritores comportamentais durante as amostras. A Figura 20 ilustra o processo de obtenção dos atributos estatísticos dos descritores comportamentais. Na primeira etapa foram extraídos os descritores comportamentais de cada quadro dos segmentos de vídeo de cada categoria comportamental registrada. As barras verticais representam as imagens (quadros do segmento). Os círculos representam os descritores comportamentais extraídos de cada imagem. Os atributos de estatística descritiva (*e.g.*, média, moda, variância) foram calculados para cada descritor de imagem, em relação a suas medições individuais na amostra comportamental em análise.

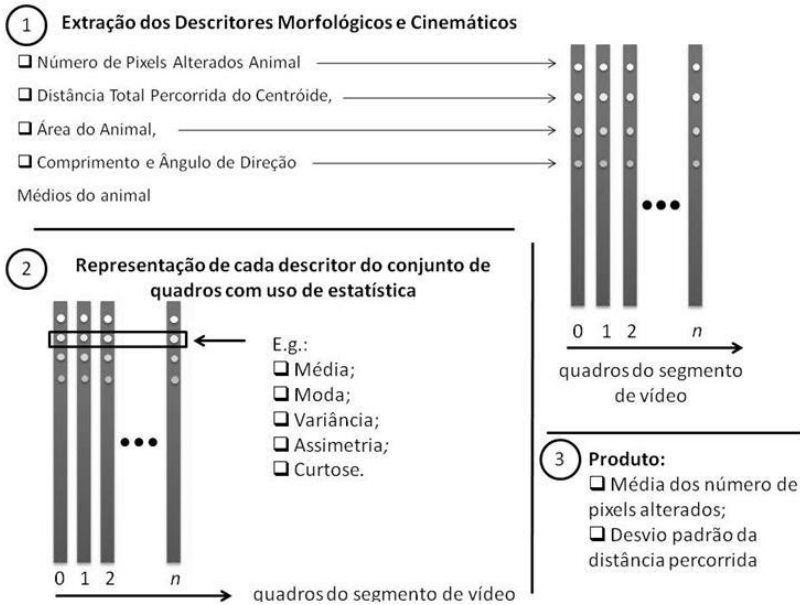


Figura 20. Representação da Informação temporal dos quadros de um comportamento

O uso de atributos estatísticos reduz o número de neurônios necessários na camada de entrada de redes neurais, pela compressão dos dados de descritores de um período de tempo. Além disso, essa compressão permite tanto o treinamento de classificadores com amostras de tamanhos distintos, quanto seu posterior uso (em modo de teste) com amostras de tamanho variável.

Os atributos estatísticos estudados foram: média, média geométrica, média harmônica, mediana, variância, desvio padrão, coeficiente de variação, erro padrão da média, assimetria e curtose (extraídos através da ferramenta STATISTICA, STATSOFT, versão 8.0). Esses atributos foram escolhidos por sugerirem diferenças estatisticamente significantes entre as categorias de comportamentos.

Os atributos de média (média aritmética), mediana, variância, e erro padrão da média são elementos comuns na descrição de variáveis estatísticas. No entanto, os atributos de média geométrica, média harmônica, coeficiente de variação, assimetria e curtose são menos comuns, e por isso serão brevemente descritos.

A assimetria mede o afastamento de uma distribuição de frequências em relação a uma distribuição simétrica. Se a assimetria é diferente de zero, a distribuição é assimétrica (*e.g.*, uma distribuição normal é perfeitamente simétrica). Complementarmente, a curtose é uma medida do achatamento de uma distribuição em relação a uma distribuição normal. E procura quantificar se uma distribuição é mais achatada ou mais pontiaguda do que a distribuição normal [88][89].

A média geométrica é usada na medição de escalas não lineares, como o cálculo de médias de razões, taxas de variações, e índices econômicos. Trata-se da n -ésima raiz do produto de n números positivos. A média harmônica é útil na análise de frequência de valores (*e.g.*, frequências musicais, variações de velocidade). Para um conjunto n de números x_i (*e.g.*, x_1, x_2, \dots, x_n), a media harmônica define-se como: n dividido pela soma dos inversos desses números. Essas duas médias foram selecionadas também pela possibilidade de descrever outras tendências dos dados de descritores (*e.g.*, componente cinemática), desconsiderados pela média aritmética. Por fim, o coeficiente de variação indica a magnitude relativa do desvio padrão em relação à média aritmética da distribuição, em forma de porcentagem [89][90].

3.5. TREINAMENTO E AVALIAÇÃO DAS REDES NEURAIS ARTIFICIAIS

Treinaram-se grupos diferentes de redes neurais, cada qual orientado a detecção de uma das categorias de interesse. Essa abordagem permitiu ajustar os procedimentos de treinamento especificamente para cada categoria comportamental. Por exemplo, o ponto de parada ótimo do treinamento para o reconhecimento do comportamento de locomoção provavelmente será distinto do mais adequado para a limpeza corporal. Além disso, os resultados da prova estatística indicavam que os parâmetros de entrada ótimos diferiam de acordo com a categoria comportamental a detectar.

Cada *perceptron* de múltiplas camadas foi projetado com três camadas. O número de neurônios da primeira camada foi determinado pelo número de parâmetros de entrada usados. Como cada rede especializou-se em apenas um comportamento, a camada de saída continha apenas um neurônio. O número de neurônios na camada oculta foi determinado empiricamente, e de forma iterativa. A partir da estrutura de neurônios que atingiu os maiores desempenho de identificação, procurou-se reduzir o número de neurônios usados na

camada intermediária, enquanto não fossem visualizados decréscimos no desempenho do método. A Tabela 2 apresenta o número de neurônios usados nas camadas de entrada e oculta das redes neurais de cada categoria de comportamento.

Tabela 2. Número de Neurônios nas Camadas das RNA-MLP de maior desempenho

Comportamentos	Número de Neurônios	
	Camada de Entrada	Camada Oculta
Locomoção	8	7
Imobilidade	8	7
Exp. Vertical	6	4
Limpeza corporal	5	3

O algoritmo de treinamento empregado foi a retropropagação do erro com taxa de momento de RUMELHART, HINTON E WILLIAMS [57]. Os parâmetros de treinamento desse algoritmo (*e.g.*, taxa de aprendizagem, constante de momento, constante de inclinação da função de ativação) também foram selecionados empiricamente (vide Tabela 3). A Figura 21 ilustra a função sigmóide logística usada nas redes neurais treinadas. O parâmetro α ilustrado altera a inclinação da curva produzida pela função, que por sua vez altera o comportamento de aprendizagem do modelo.

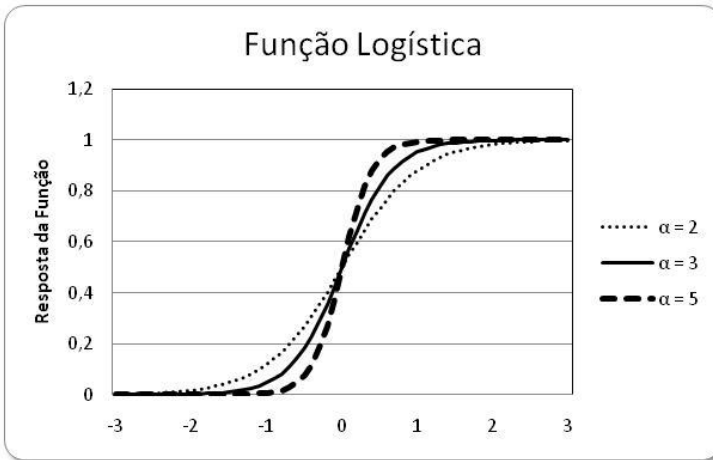


Figura 21. Exemplos de funções logísticas

Tabela 3. Parâmetros usados no algoritmo de treinamento dos MLP

Comportamento	Taxa Aprendizagem	Constante de Momento
Locomoção	0,01	0,50 a 0,70
Imobilidade	0,01	0,50 a 0,60
Exploração Vertical	0,01 a 0,10	0,30 a 0,70
Limpeza Corporal	0,01 a 0,10	0,40

O procedimento de validação cruzada foi empregado para avaliar a generalização do modelo, e os índices de sensibilidade, especificidade, AUC, eficiência, e coeficiente Kappa para avaliar o desempenho de registro dos eventos comportamentais. A validação cruzada múltipla (ou de n -dobras) foi usada.

Essa variante de validação cruzada consiste em dividir o conjunto de exemplos em n subconjuntos. O treinamento da rede neural foi repetido dez vezes, e a cada repetição um dos subconjuntos é usado como grupo de validação, e os outros nove são agrupados e usados como conjunto de treinamento [19]. A distribuição das amostras entre os subconjunto (ou dobras) da validação cruzada foi realizada de forma automatizada e aleatória. Cada dobra possuiu a mesma quantidade de exemplos de cada categoria de comportamentos, sendo que nenhuma amostra repetiu-se entre as dobras [16][91]. Apesar de aumentar o esforço computacional empregado, quando comparado ao uso de apenas um conjunto de treino e validação, o uso da validação múltipla permite detectar tendências de resultados devido à distribuição das amostras entre os conjuntos de treinamento e validação.

O procedimento de avaliação de desempenho foi executado de forma integrada ao procedimento de treinamento das RNAs [64]. Ao fim de cada época de treinamento, congelavam-se os pesos da RNA, e além de medir os erros médios quadráticos sobre o conjunto de treino e o de validação, mediam-se os índices de desempenho da rede no reconhecimento dos eventos comportamentais sobre o conjunto de validação.

Nas próximas seções serão descritos o método de pré-seleção usado para determinar a relevância dos atributos de descritores para o procedimento de descrição e identificação das categorias comportamentais (usando as RNA-MLP), e os testes que avaliaram a capacidade de identificação dos classificadores produzidos em situações em que os padrões comportamentais estão alterados.

3.6. VALIDAÇÃO DA PRÉ-SELEÇÃO DOS DESCRITORES COMPORTAMENTAIS

A adição de um conjunto de atributos estatísticos para cada descritor comportamental implicou num razoável aumento da quantidade de parâmetros-candidatos às entradas de modelos de classificação do comportamento. Parece importante pré-selecionar (ou avaliar a relevância) de cada parâmetro de entrada antes de o mesmo compor as entradas de modelos de redes neurais artificiais (2º objetivo específico).

Como haviam sido usados atributos estatísticos para descrever a variação dos descritores, optou-se por avaliar/validar o uso de provas estatísticas de hipótese como procedimentos de pré-seleção de parâmetros, devido a sua aplicabilidade em outros domínios de problema [73][76][77][92].

Existem duas classes de provas estatísticas: as provas paramétricas e as não-paramétricas. Uma prova estatística paramétrica é aquela cujo modelo especifica certas condições que os parâmetros da população da qual se extraíram as amostras devem atender. A validade dos resultados obtidos na prova paramétrica depende diretamente da validade dessas suposições sobre os dados. Uma prova não-paramétrica não especifica condições sobre os parâmetros da população da qual se extraíram as amostras, apenas algumas suposições básicas sobre os dados (*e.g.*, observações independentes) [2][5][93], e por isso são conhecidas como provas livres de distribuição.

Os resultados de provas paramétricas têm maior significância (ou validade), pois especificam certas condições sobre os dados, as quais reduzem a probabilidade desses testes cometerem erros, *e.g.*, rejeitarem uma hipótese quando ela é verdadeira. Entre as condições de uso que devem ser cumpridas, estão: os dados devem advir de uma população com distribuição normal, e as variâncias entre as medições de um parâmetro dos grupos em teste (*e.g.*, locomoção e exploração vertical) devem ser iguais (variâncias homocedásticas). O uso dessas provas quando os dados não cumprem as condições especificadas reduz a validade de seus resultados [2][5][93].

Portanto, a seleção da prova estatística para pré-seleção dos parâmetros de entrada do modelo de RNA iniciou com o estudo da distribuição das frequências dos atributos estatísticos dos descritores. As amostras de comportamentos foram agrupadas independentemente dos ratos de que advinham, pois se procurou selecionar descritores que

fossem independentes das diferenças existentes entre animais de uma mesma espécie.

Resumidamente, iniciou-se com testes de normalidade da distribuição das amostras de dados (de cada atributo estatístico de cada descritor), usando o teste de Shapiro-Wilks [94] (de maior robustez em relação aos testes de Kolmogorov-Smirnov e de Lilliefors) [95][96]. Em seguida, avaliou-se a homocedasticidade das variâncias entre grupos comportamentais dos parâmetros oriundos de amostras com distribuição normal.

No teste de Shapiro-Wilk a hipótese nula (h_0) afirma que as amostras advêm de uma distribuição normal, e a hipótese alternativa (h_a) afirma que os dados advêm de outro tipo de distribuição (não-normal). Ou seja, quando o índice de significância (p) é significativo (*e.g.*, $p < 0,05$), conclui-se que os dados em estudo advêm de uma distribuição não-normal. Os atributos de assimetria da área e de médias geométrica e harmônica do comprimento do animal foram os únicos indicados como normalmente distribuídos. Uma inspeção visual dos histogramas da Figura 22 confirma esses resultados.

O teste de homocedasticidade usado foi o de Brown-Forsythe

. Previamente ao emprego desse método foram retiradas todas as amostras de cada par atributo/descritor que foram consideradas *outliers*. Foram marcados como *outliers* os elementos cujo valor fosse maior do que o valor da média somado a duas unidades de desvio padrão, ou que seu valor fosse menor do que o valor da média subtraído de duas unidades de desvio padrão. A média e o desvio padrão de referência foram calculados para cada par atributo estatístico/descritor comportamental de cada comportamento, sobre as amostras de todos os animais. Apenas o grupo de dados da assimetria da área apresentou homocedasticidade entre as variâncias (p de 0,0971). Os grupos de média geométrica e de média harmônica foram considerados heterocedásticos (variâncias diferentes, p de 0,001).

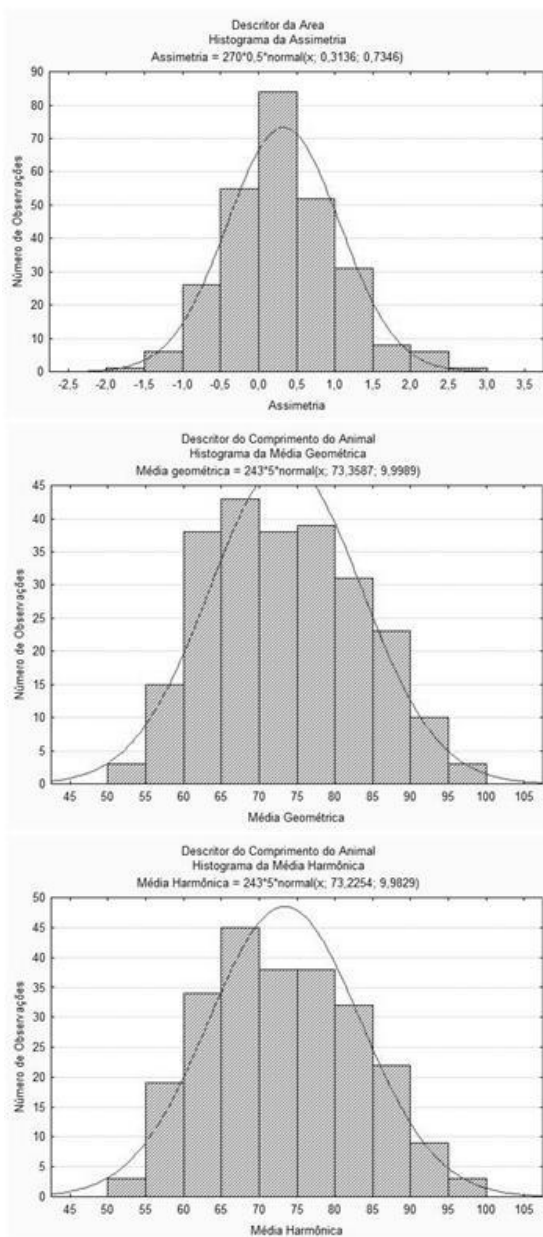


Figura 22. Histogramas dos atributos de assimetria da área e de médias geométrica e harmônica do comprimento do animal

Como apenas a área da assimetria atendeu as condições necessárias para o uso de testes paramétricos, padronizou-se o uso de testes não-paramétricos para avaliação de todos os casos. Escolheu-se a análise de variância de Kruskal-Wallis como prova estatística. No entanto, essa prova apenas identifica se existem diferenças entre os grupos de comportamentos, e não indica quais grupos diferem entre si. Nos casos em que foram encontradas diferenças, os grupos comportamentais do par atributo/descritor em análise foram agrupados em pares, e o teste de Mann-Whitney usado para verificar quais grupos desse par diferenciava significativamente.

Foram selecionados os pares de atributo e descritor que indicavam diferenças significativas entre a categoria comportamental em análise e todas as (se não a maioria das) outras categorias em estudo. Elaborou-se um procedimento de validação para testar se os resultados obtidos eram relevantes para o treinamento de RNAs.

O procedimento de validação consistiu em treinar redes neurais artificiais a partir de parâmetros de entrada (índices comportamentais) pré-selecionados pela prova estatística (Grupo Pré-seleção). Os resultados do grupo pré-selecionado foram comparados com o desempenho de outros dois grupos que serviram de controle: 1) um grupo usando apenas os atributos de média e de desvio padrão de cada descritor comportamental; 2) e um grupo usando índices que não haviam sido pré-selecionados pela prova estatística.

O procedimento de validação foi realizado para cada comportamento de interesse, pois a prova estatística pré-selecionou diferentes índices de maior significância em cada categoria de comportamento. O procedimento de treinamento das RNA está descrito em detalhes na seção 3.5. Discutem-se a seguir os detalhes de cada grupo de validação.

3.6.1. Grupo de atributos pré-selecionados

Este grupo abrange os índices comportamentais pré-selecionados pelas provas estatísticas de Kruskal-Wallis e Mann-Whitney. Minimizou-se o número atributos estatísticos a dois elementos para cada descritor comportamental usado. A escolha dos dois atributos privilegiou o uso de uma medida de tendência (*e.g.*, média, mediana) e uma medida de dispersão (*e.g.*, desvio padrão, erro padrão), quando possível. O critério de desempate entre duas medidas de mesma

categoria foi o nível de significância obtido no teste de Mann-Whitney. Ou seja, foram selecionados os índices com menor índice de significância. Quando os níveis de significância eram iguais ou inferiores a 10^{-5} , optou-se pelo atributo estatístico mais comum entre os descritores morfológicos e cinemáticos do comportamento em análise. A Quadro 4 ilustra o grupo de índices pré-selecionados.

Os descritores ângulo e variância angular foram removidos desse grupo, pois independente da categoria em análise, nenhum dos atributos estatísticos extraídos desses descritores era capaz de diferenciar a categoria em análise das outras categorias de interesse, apenas diferenciações parciais entre algumas categorias. Por exemplo, segundo a prova estatística, os atributos de variação angular diferenciavam as amostras de locomoção de imobilidade, porém não eram capazes de diferenciar locomoção de exploração vertical.

Descritores usados	Comportamentos			
	Locomoção	Imobilidade	Exploração Vertical	Limpeza corporal
Área do Animal	Mediana, Coeficiente de variação	Mediana, e Erro padrão da média	Desvio padrão	-
Comprimento do animal	Mediana, Desvio padrão	Média, e Erro padrão da média	Desvio padrão	Erro padrão da média
Distância Percorrida	Mediana, e Desvio padrão	Média, e Erro padrão da média	Média, e Desvio padrão	Média, e Desvio padrão
Número de pixels alterados	Mediana, e Desvio padrão	Mediana, e Erro padrão da média	Média, e Erro padrão da média	Mediana, e Erro padrão da média

Quadro 4. Atributos estatísticos de cada descritor morfológico/cinemáticos selecionados pela prova estatística para a diferenciação entre as categorias comportamentais

Nos casos em que a média e a mediana obtiveram os mesmos níveis de significância, a mediana foi preferida para padronizar os atributos usados, já que em vários casos esse atributo foi pré-selecionado por exibir maior nível de significância que a média na diferenciação estatística entre as categorias de comportamentos.

3.6.2. Grupos de atributos de controle

Foram usados dois grupos de controle para teste da relevância dos atributos selecionados como parâmetros de entrada das redes neurais. O primeiro grupo de controle a média e o desvio padrão de cada descritor comportamental, enquanto o segundo grupo usou atributos estatísticos que não foram pré-selecionados em nenhuma etapa pela prova estatística.

O grupo de atributos composto da média e do desvio padrão de cada descritor de imagem avaliou se havia diferenças em desempenho entre os classificadores que usavam atributos pré-selecionados, e classificadores treinados a partir de uma escolha padronizada de atributos. A Figura 23 ilustra os parâmetros de média e desvio padrão dos descritores comportamentais (*e.g.*, distância percorrida, número de pixels alterados) por categoria de comportamento. Em geral, todos os atributos apresentados indicam diferenças significantes entre as categorias de comportamentos (ao teste de Kruskal-Wallis, $p < 0,05$). Uma exceção foi a média do comprimento do animal, que foi incapaz de diferenciar as categorias de exploração vertical de limpeza corporal.

O segundo grupo de controle continha os atributos estatísticos não selecionados pela prova de Kruskal-Wallis. O atributo de assimetria da área do animal e os atributos de média geométrica e média harmônica do comprimento do animal foram retirados, pois tinham sido selecionados pela prova paramétrica testada. O Quadro 5 ilustra os descritores usados nesse grupo de controle por categoria de comportamento. Atributos de média, mediana, e desvio padrão foram preferencialmente incluídos nesse grupo quando não faziam parte dos pré-selecionados pela prova, para dessa forma confirmar a sua irrelevância para o processo de descrição do comportamento.

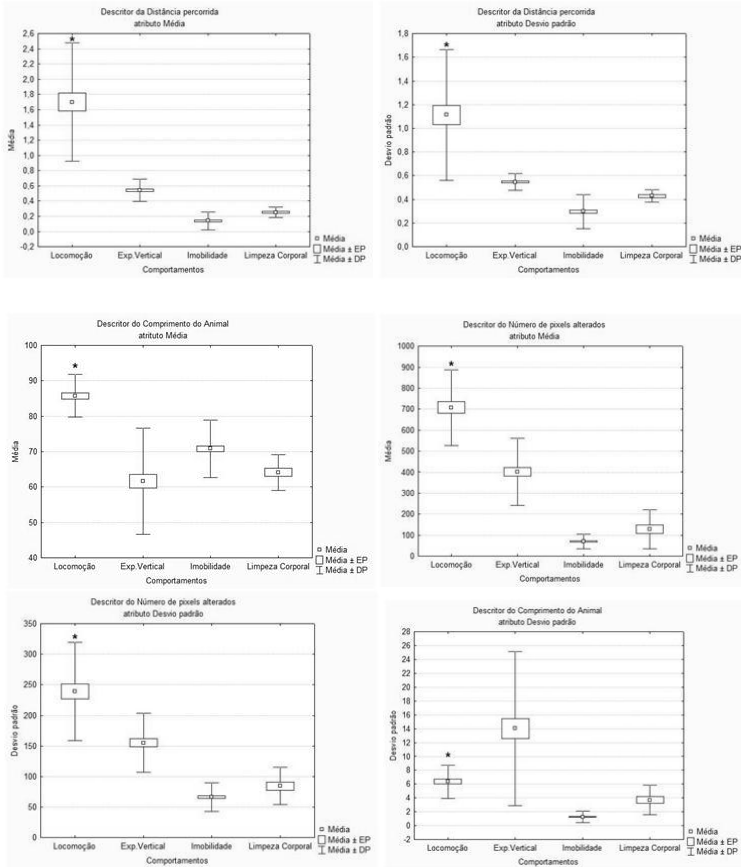


Figura 23. Atributos de média e desvio padrão dos descritores comportamentais (distância percorrida, número de pixels alterados, e comprimento do animal) por categoria comportamental. (*) indica diferenças estatisticamente significativas entre a locomoção e as outras categorias analisadas (Kruskal-Wallis, $p < 0.05$).

Descritores usados	Comportamentos			
	Locomoção	Imobilidade	Exploração Vertical	Limpeza Corporal
Ângulo	Média, Coeficiente de variação	Mediana, Assimetria	Assimetria	-
Área	Curtose	Média, Curtose	Mediana	Média
Comprimento do animal	Curtose	Assimetria	Mediana	Assimetria
Distância percorrida	Assimetria	Mediana	Curtose	Assimetria
Número de pixels alterados	Curtose	Média geométrica	Curtose	Curtose
Variação Angular	Média, Desvio padrão	Assimetria	Assimetria	Média

Quadro 5. Grupo de Controle II - Complemento da Pré-seleção

3.7. AVALIAÇÃO DOS DESCRITORES

Os descritores comportamentais e os respectivos atributos estatísticos foram testados em sua capacidade de identificar os comportamentos em situações nas quais os padrões comportamentais estão em sua forma “padrão”, e quando alterados pelo uso de tratamentos.

O treinamento das RNAs usou apenas amostras de comportamentos de animais ingênuos a tratamentos. As épocas de maior desempenho na identificação de cada comportamento foram selecionadas como indicadores do desempenho do método. Conforme discutido na seção 3.6, foram testadas diferentes combinações de atributos e descritores para avaliar e validar o método de pré-seleção de parâmetros de entrada empregado (prova de Kruskal-Wallis).

Após a validação, os classificadores de maior desempenho de cada conjunto de atributos avaliados quanto seu desempenho na identificação de comportamentos alterados pelo uso de tratamento. Os classificadores operaram em modo de teste, portanto não houve ajustes nos parâmetros internos dessas redes.

No primeiro teste avaliou-se o desempenho de identificação dos classificadores frente a amostras comportamentais extraídas de vídeos de animais tratados com cafeína (conhecida por alterar a atividade locomotora de roedores). A expressão padrão do comportamento é comumente alterada por estudos de disciplinas de neurociências e farmacologia, como forma de avaliar os efeitos de drogas sobre o organismo (vide seção 2.2). Apesar de essas alterações serem quantificadas e avaliadas em relação a modificações nos padrões de frequência e duração de cada categoria comportamental, suspeitava-se que as drogas também afetavam os padrões comportamentais extraídos pelos descritores. Das duas doses de tratamento testadas (vide protocolo experimental, 3.1.4), os comportamentos de animais tratados com cafeína na dose de 6 mg/kg apresentaram diferenças estatisticamente significativas ($p < 0,05$) na comparação com os animais sem tratamento. Essas diferenças referiam-se a alterações nos padrões de frequência e duração do comportamento de exploração vertical, e na distância percorrida pelos animais.

O segundo teste avaliou a capacidade dos classificadores em identificar comportamentos em um segmento de vídeo de tamanho fixo. Os segmentos de vídeo usados para o treinamento das redes possuíam durações de tempo variadas, e se referiam ao período de tempo exato de uma amostra comportamental. O uso de um segmento de tamanho fixo avalia a capacidade do classificador em identificar um segmento de vídeo tanto completo como parcial do comportamento. Escolheu-se a duração de 1 segundo, pois era a menor duração encontrada entre as amostras comportamentais usadas neste trabalho (veja Figura 24).

Foram selecionadas aleatoriamente duas amostras de cada categoria comportamental por rato usado no protocolo experimental (Total: 31 amostras). Essas amostras foram retiradas do conjunto de amostras usadas para treinamento e avaliação das redes (ratos ingênuos ao tratamento). As amostras comportamentais selecionadas de duração superior a definida foram subamostradas. A subamostragem consistiu da delimitação do quadro central do período de tempo da amostra (número de quadros da amostra dividido por 2), e a partir dessa posição central, obteve-se os descritores dos 15 quadros que o precediam, do próprio quadro, e dos 14 quadros que o sucediam.

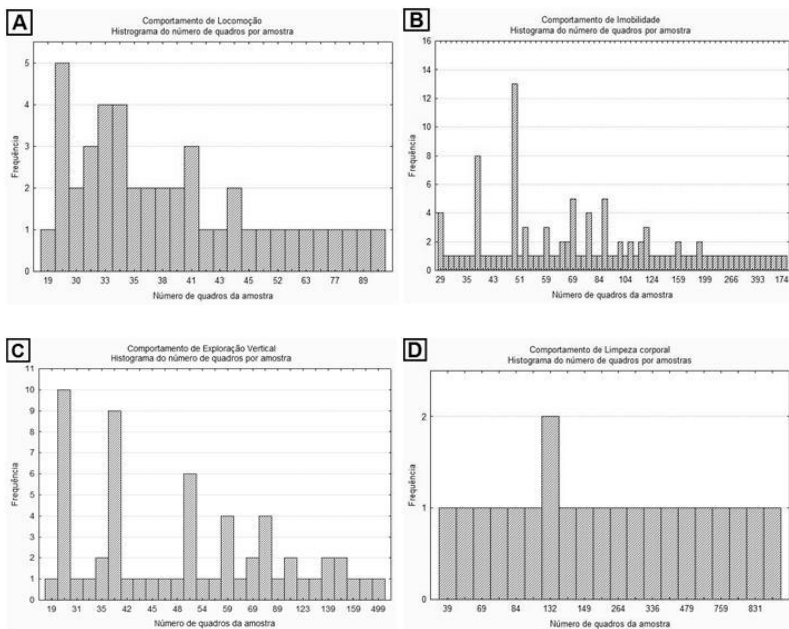


Figura 24. Histogramas do número de quadros das amostras comportamentais das categorias comportamentais de locomoção (A), imobildade (B), exploração vertical (C), e limpeza corporal (D).

4. RESULTADOS

4.1. INTRODUÇÃO

Neste capítulo são descritos os resultados obtidos com o emprego de descritores comportamentais (morfológicos e cinemáticos) como descritores das imagens comportamentais no reconhecimento automático de comportamentos através de redes neurais artificiais. Os resultados estão agrupados de acordo com a análise executada: validação do procedimento de pré-seleção dos parâmetros de entrada (4.2), identificação de comportamentos (4.3), e a análise dos índices de desempenho (4.4). Os resultados expressos em termos de média e desvio padrão referem-se ao desempenho médio dos classificadores de maior desempenho nas 10 dobras da validação cruzada sob o conjunto de validação ou em novos conjuntos de teste.

4.2. PRÉ-SELEÇÃO DE PARÂMETROS

As Figuras 25, 26, 27 e 28 ilustram o índice AUC do classificador de maior desempenho de identificação de cada categoria comportamental no conjunto de validação, de cada uma das 10 dobras da validação cruzada estratificada. Os resultados estão subdivididos por conjunto de atributos usado para treinar os classificadores.

A Figura 25 ilustra o desempenho dos classificadores treinados para identificar a locomoção. Visualizam-se dobras do conjunto de atributos não selecionados que não foram capazes de identificar esse comportamento (desempenho de zero). Verificam-se diferenças significativas ($p < 0,05$) entre o desempenho dos classificadores treinados a partir dos atributos selecionados e aqueles treinados a partir dos não selecionados (ao teste de Mann-Whitney).

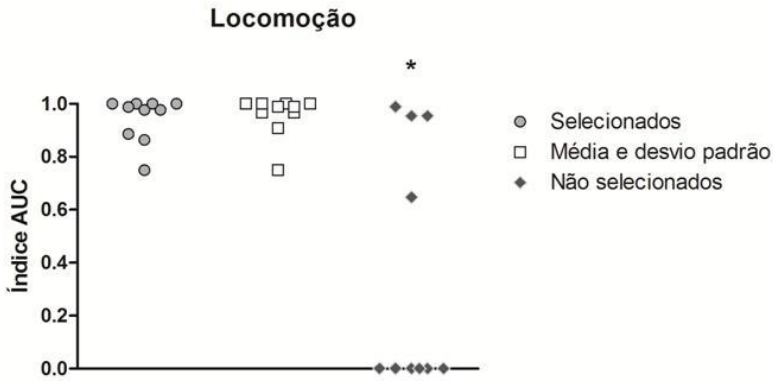


Figura 25. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de Locomoção, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney).

A Figura 26 ilustra o desempenho dos classificadores treinados para identificar a imobilidade. Como na locomoção, visualizam-se dobras do conjunto de atributos não selecionados que não foram capazes de identificar o comportamento (desempenho de zero). Também existem diferenças significativas ($p < 0,05$) entre o desempenho dos classificadores treinados a partir dos atributos selecionados e aqueles treinados a partir dos não selecionados (ao teste de Mann-Whitney).

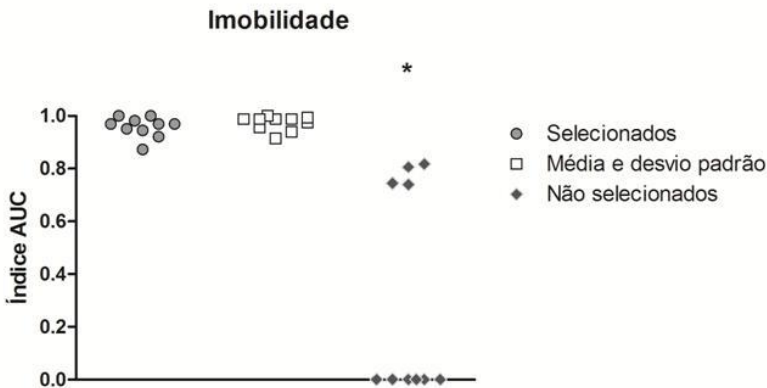


Figura 26. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de Imobilidade, por grupo de atributos usado

para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney)

A Figura 27 ilustra o desempenho dos classificadores treinados para identificar a limpeza corporal. Neste caso visualizam-se apenas uma dobra do conjunto de atributos não selecionados que teve desempenho acima de 50%. Verificam-se diferenças significantes ($p < 0,05$) entre o desempenho dos classificadores treinados a partir dos atributos selecionados e aqueles treinados a partir dos não selecionados (ao teste de Mann-Whitney).

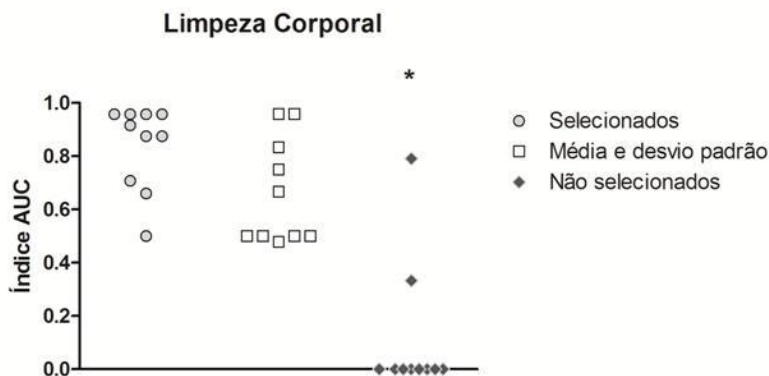


Figura 27. Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de Limpeza Corporal, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney).

A Figura 28 ilustra o desempenho dos classificadores treinados para identificar a exploração vertical. Visualizam-se apenas três dobras do conjunto de atributos não selecionados que não foram capazes de identificar esse comportamento. Verificam-se diferenças significantes ($p < 0,05$) entre o desempenho dos classificadores treinados a partir dos atributos selecionados e aqueles treinados a partir dos não selecionados (ao teste de Mann-Whitney).

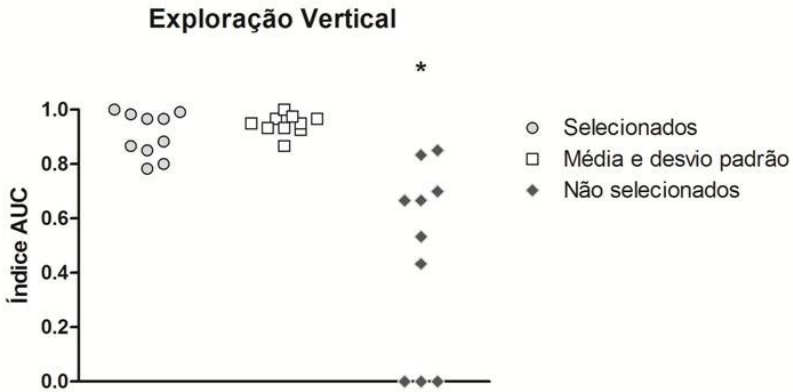


Figura 28. . Índices AUC dos classificadores nas 10 dobras da validação cruzada estratificada do comportamento de exploração vertical, por grupo de atributos usado para treinar as RNAs. (*) $p < 0,05$ indica diferenças significantes em comparação com o grupo de atributos selecionados (Teste de Mann-Whitney).

4.3. IDENTIFICAÇÃO DOS COMPORTAMENTOS

Dos grupos de atributos testados na validação do procedimento de pré-seleção, selecionaram-se os classificadores dos grupos de atributos que apresentaram o maior desempenho médio de identificação em cada categoria comportamental. A Tabela 4 apresenta o desempenho médio de identificação dos classificadores em cada categoria comportamental. O desempenho foi medido usando três índices de desempenho distintos.

Tabela 4. Nível de desempenho de identificação dos comportamentos segundo índices de eficiência, AUC, e Kappa

Comportamentos	Conjunto de atributos	Índice de desempenho (média ± desvio padrão)		
		Eficiência (%)	Kappa (%)	AUC (%)
Locomoção	Média e desvio padrão	97,3 ± 2,0	90,9 ± 8,0	95,6 ± 8,0
Imobilidade	Média e desvio padrão	95,7 ± 2,0	91,8 ± 4,0	97,3 ± 2,0
Exploração Vertical	Média e desvio padrão	94,6 ± 3,0	86,4 ± 7,0	94,6 ± 3,0
Limpeza corporal	Pré-seleção	96,1 ± 0,0	75,4 ± 2,0	83,6 ± 16,0

Os resultados apresentados referem-se ao uso de amostras de comportamentos de animais ingênuos a tratamentos com drogas. Usando-se de amostras de comportamentos extraídas de vídeos experimentais dos mesmos animais, porém agora tratados com cafeína, avaliou-se o desempenho de identificação dos classificadores em cada categoria comportamental. Porém, restringiu-se apenas aos classificadores treinados a partir dos atributos pré-selecionados e daqueles treinados usando a média e o desvio padrão dos descritores (classificadores de maior desempenho na validação da pré-seleção, vide Figuras 26 e 27). Para verificar se o desempenho dos dois conjuntos de atributos realmente diferia na identificação de uma mesma categoria comportamental, avaliou-se o desempenho obtido no índice AUC quanto a diferenças significativas usando o teste de Mann-Whitney.

A Tabela 5 apresenta a média e o desvio padrão do desempenho de identificação dos 10 classificadores da validação cruzada de cada categoria comportamental na identificação das amostras de animais tratados com cafeína 2 e 6 mg/kg. Os resultados em cada cenário são apresentados para ambos os conjunto de atributos usados para treinar os classificadores.

Tabela 5. Desempenho dos classificadores na validação cruzada de 10 dobras usando amostras de comportamentos extraídas de vídeos de animais tratados com cafeína

	Dose de cafeína	Conjunto de atributos	Índices de Desempenho (média ± desvio padrão, %)		
			Eficiência	Kappa	AUC
Locomoção	2 mg/kg	Pré-seleção	92,2 ± 0,5	75,6± 1,2	93,4 ± 1,1
		Média e desvio padrão	96,7 ± 1,5	88,1± 4,3	97,5 ± 2,6
	6 mg/kg	Pré-seleção	94,0 ± 0,7	78,4± 2,0	83,5 ± 0,9
		Média e desvio padrão	93,8 ± 0,6	77,4± 1,8	83,8 ± 1,7
Imobilidade	2 mg/kg	Pré-seleção	95,6 ± 0,3	85,8 ± 0,7	97,8 ± 0,6
		Média e desvio padrão	96,2 ± 0,2	87,7± 0,6	95,6 ± 0,4
	6 mg/kg	Pré-seleção	82,8 ± 1,0	69,1± 0,9	85,2 ± 1,8
		Média e desvio padrão	81,5 ± 0,1	67,2± 0,1	78,1 ± 1,1
Exploração Vertical	2 mg/kg	Pré-seleção	81,9 ± 2,0	61,9± 2,5	71,8 ± 3,0
		Média e desvio padrão	77,3 ± 2,0	56,1± 3,4	68,9 ± 4,0
	6 mg/kg	Pré-seleção	77,4 ± 8,0	56,7± 2,0	67,0 ± 2,0
		Média e desvio padrão	80,3 ± 2,0	61,3± 3,1	65,1 ± 7,0
Limpeza Corporal	2 mg/kg	Pré-seleção	93,7 ± 0,2	50,0± 2,1	54,2 ± 8,0
		Média e desvio padrão	93,6 ± 0,0	48,3± 0,0	51,0 ± 7,0
	6 mg/kg	Pré-seleção	98,5 ± 0,1	61,9± 2,9	78,0± 10,0
		Média e desvio padrão	98,2 ± 0,0	49,4± 0,3	50,4± 10,0

A Tabela 6 ilustra as diferenças estatística de desempenho dos classificadores entre identificando comportamentos do conjunto de validação (animais ingênuos a tratamentos), e o desempenho de identificação em amostras de animais tratados com cafeína 6 mg/kg.

Tabela 6. Nível de significância das diferenças estatísticas de desempenho entre a identificação de comportamentos de animais tratados com cafeína 6mg/kg em relação à identificação de amostras de comportamentos de animais sem tratamento

Comportamentos	Grupo de atributos usado nos classificadores	
	Pré-selecionados	Média e desvio padrão
Locomoção	0,0200	0,0020
Imobilidade	0,0001	0,0001
Exploração Vertical	0,0001	0,0001
Limpeza Corporal	0,0680	0,0800

A Tabela 7 apresenta o nível de significância obtido no teste de Mann-Whitney na avaliação de diferenças estatisticamente significantes entre o desempenho dos classificadores usando os atributos pré-selecionados e aqueles usando apenas a média e o desvio padrão dos descritores.

Tabela 7. Nível de significância das diferenças estatísticas de desempenho entre os classificadores usando grupos de atributos pré-selecionados, e aqueles usando apenas a média e o desvio padrão, na identificação de comportamentos em amostras de ratos tratados com cafeína 2 e 6 mg/kg.

Comportamentos	Dose de cafeína	Nível de significância no teste de Mann-Whitney
Locomoção	2 mg/kg	0,0002
	6 mg/kg	Não significante
Imobilidade	2 mg/kg	0,0001
	6 mg/kg	0,0001
Exploração Vertical	2 mg/kg	0,0900
	6 mg/kg	0,1100
Limpeza Corporal	2 mg/kg	0,0400
	6 mg/kg	0,0001

A avaliação dos classificadores quanto à influência do número de quadros usados para calcular os atributos estatísticos sobre o desempenho de identificação dos classificadores usou segmentos de 1 segundo dos comportamentos de animais não tratados. A Tabela 8 apresenta o desempenho dos classificadores nesse cenário, por categoria comportamental (média \pm desvio padrão).

Tabela 8. Desempenho dos classificadores na validação cruzada de 10-dobras usando amostras de 30 quadros (≈ 1 segundo)

	Atributos de Descritores	Índices de Desempenho (média \pm desvio padrão, %)		
		Eficiência	Kappa	AUC
Locomoção	Pré-seleção	82,57 \pm 0,04	63,58 \pm 7,74	76,24 \pm 0,06
	Média e desvio padrão	82,25 \pm 0,05	62,01 \pm 9,82	59,91 \pm 0,23
Imobilidade	Pré-seleção	88,06 \pm 0,02	73,90 \pm 4,34	88,80 \pm 0,02
	Média e desvio padrão	87,40 \pm 0,01	71,54 \pm 2,70	88,36 \pm 0,04
Exploração Vertical	Pré-seleção	74,83 \pm 0,01	45,03 \pm 3,93	24,37 \pm 0,14
	Média e desvio padrão	79,35 \pm 0,01	60,31 \pm 2,42	68,96 \pm 0,03
Limpeza Corporal	Pré-seleção	77,41 \pm 0,00	47,18 \pm 2,48	51,60 \pm 0,09
	Média e desvio padrão	78,05 \pm 1,35	45,99 \pm 0,04	38,68 \pm 0,17

A Figura 29 ilustra a sequência de comportamentos analisada pelo classificador, e o resultado obtido para cada comportamento. Para os comportamentos de exploração vertical e imobilidade são exibidos dois classificadores, pois os índices de desempenho usados discordaram no classificador escolhido de maior desempenho entre as 10-dobras da validação cruzada.

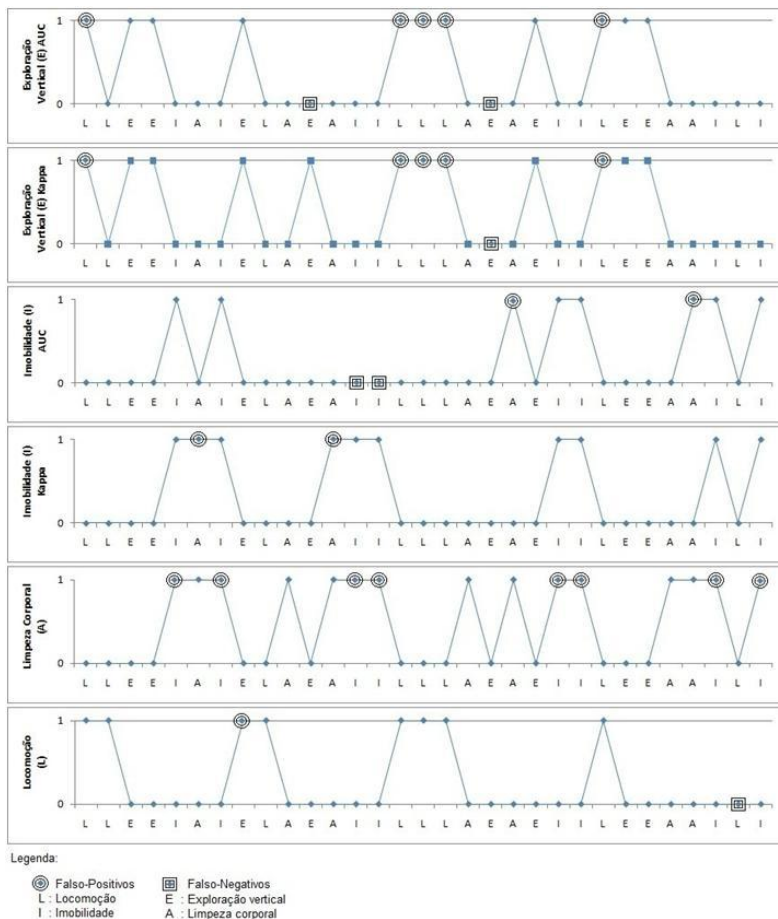


Figura 29. Desempenho de detecção dos classificadores em amostras de 1 segundo (30 quadros) dos eventos comportamentais. O eixo vertical apresenta a resposta do classificador (1 – comportamento detectado, 0 – outro evento). O eixo horizontal apresenta a sequência de categorias de comportamentos usados no teste do classificador (L – Locomoção, I – Imobilidade, E – Exploração Vertical, A – Limpeza Corporal). Para os comportamentos de exploração vertical e imobilidade são ilustrados dois classificadores, pois os índices AUC e Kappa discordaram no classificador de maior desempenho de detecção.

A Tabela 9 descreve as diferenças estatísticas existentes entre os desempenhos dos classificadores usando atributos pré-selecionados, e aqueles usando apenas a média e o desvio padrão, na identificação de

comportamentos em segmentos de vídeo de 1 segundo de duração (de acordo com índice AUC ao teste de Mann-Whitney).

Tabela 9. Nível de significância das diferenças estatísticas de desempenho entre os classificadores usando atributos pré-selecionados e aqueles usando apenas a média e o desvio padrão dos descritores, na identificação de comportamentos em segmentos de vídeo de 1 segundo de duração.

Comportamentos	Nível de significância no teste de Mann-Whitney
Locomoção	$p < 0,0200$
Imobilidade	$p < 0,8400$
Exploração Vertical	$p < 0,0001$
Limpeza Corporal	Não significante

A Tabela 10 apresenta o nível de significância das diferenças encontradas entre o desempenho dos classificadores na identificação de comportamentos no conjunto de validação comparado ao desempenho na identificação de comportamentos em segmentos de 1 segundo (de acordo com o índice AUC e o teste de Mann-Whitney).

Tabela 10. Nível de significância do teste de Mann-Whitney na avaliação de diferenças estatísticas de desempenho dos classificadores na identificação nos comportamentos do conjunto de validação e nos comportamentos com 1 segundo de duração

Comportamentos	Grupo de atributos usado pelos classificadores	
	Pré-selecionados	Média e desvio padrão
Locomoção	0,0008	0,0002
Imobilidade	0,0010	0,0001
Exploração Vertical	0,0001	0,0001
Limpeza Corporal	0,0001	0,0220

4.4. AVALIAÇÃO DOS ÍNDICES DE DESEMPENHO

Após a avaliação do desempenho de identificação dos comportamentos, e a validação do método de pré-seleção dos parâmetros de entrada desses modelos, avaliaram-se os índices de desempenho de diagnóstico e análise de concordância quanto a representação do desempenho real das RNA treinadas. A Figura 30 ilustra o desempenho dos classificadores na validação cruzada sobre o conjunto de validação, de acordo com os índices AUC, Kappa e

eficiência. Os resultados são apresentados para os três conjuntos de atributos usados na validação do método de pré-seleção de parâmetros.

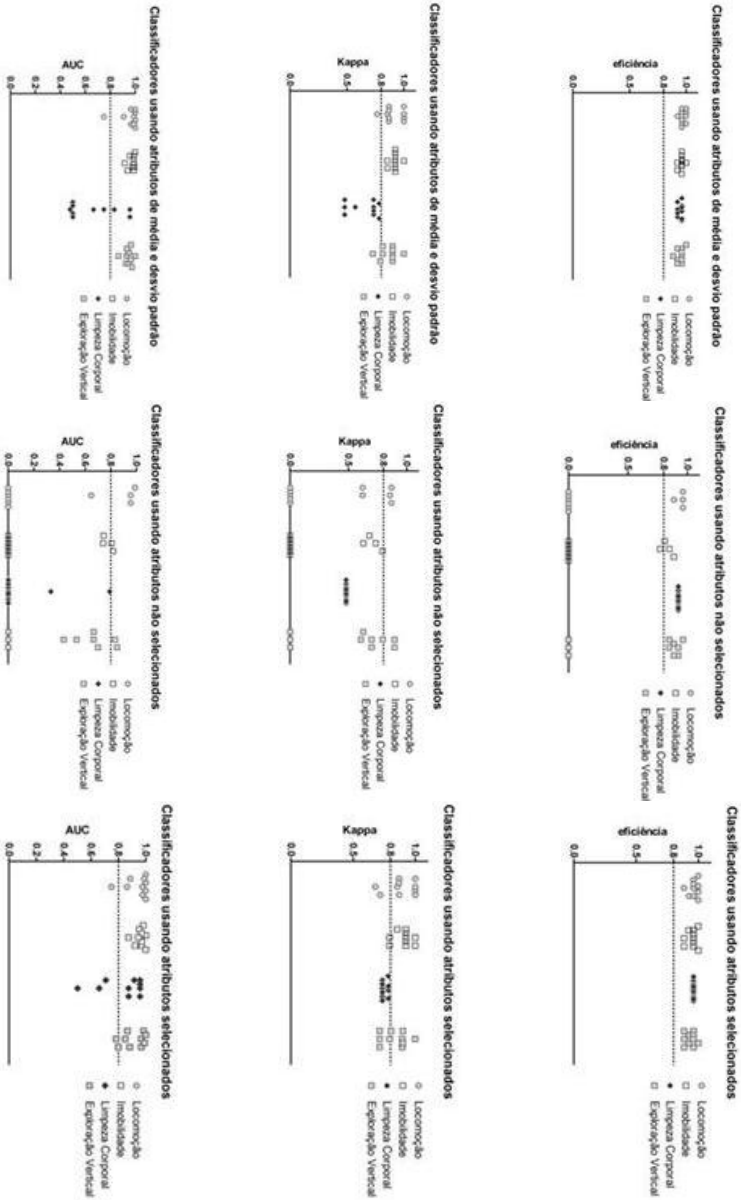


Figura 30. Desempenho de identificação comportamental dos classificadores nas 10-dobras da validação cruzada segundo diferentes índices de desempenho, por grupo de atributos usado no treinamento.

A Figura 31 ilustra o desempenho de identificação dos classificadores de limpeza corporal, segundo diferentes índices de representação do desempenho (AUC, eficiência, Kappa, Sensibilidade e Especificidade). Esses resultados referem-se aos classificadores treinados a partir do conjunto de atributos considerados irrelevantes para diferenciação desse comportamento das demais categorias (ao teste de Kruskal-Wallis).



Figura 31. Comparação de diferentes índices de desempenho na avaliação da capacidade de identificação de comportamentos de limpeza corporal por classificadores treinados a partir de atributos não selecionados pela prova estatística (segundo os índices AUC, eficiência, Kappa, sensibilidade e especificidade).

5. DISCUSSÃO

O principal objetivo desta tese foi o desenvolvimento e o teste de índices que descrevessem os padrões comportamentais contidos na morfologia da área do corpo de um animal, e na cinemática de suas ações (descritores comportamentais), para serem usados como parâmetros de entrada de Redes Neurais Artificiais.

Os descritores comportamentais foram extraídos das imagens contidas em segmentos de vídeos experimentais, através da ferramenta ETHOWATCHER®. Esta ferramenta foi desenvolvida durante esta tese, e permite tanto o registro dos comportamentos dos animais, quanto a análise de sua atividade no ambiente experimental. Por unificar essas duas funções em uma única ferramenta, o ETHOWATCHER® disponibiliza o acesso aos padrões morfológicos e cinemáticos de cada imagem do segmento de vídeo analisado. O acesso aos dados individuais de cada imagem foi fundamental no estudo e na seleção de atributos estatísticos relevantes na descrição das categorias comportamentais.

O ETHOWATCHER® está sendo distribuído gratuitamente pela internet (em www.ethowatcher.ufsc.br), como uma ferramenta para registro e análise da atividade comportamental. Já foram contabilizados 170 cadastros em junho/2011 no sistema de distribuição do software desde o lançamento do *website* em outubro de 2010 (usuários distribuídos entre os estados do Brasil, e países da América Latina, Europa, e Oriente Médio). Previamente ao seu lançamento na web, esse software já vinha sendo usado em projetos de pesquisa em diferentes centros de pesquisa da UFSC (*e.g.*, farmacologia, neurociência), no registro e na análise da atividade de roedores e peixes.

Espera-se que futuramente pesquisadores externos à instituição ampliem o uso do ETHOWATCHER® para além das funções de Etografia e análise de atividade, e comecem a contribuir diretamente com o software, desenvolvendo ferramentas que atuem sobre os dados já extraídos pelo sistema. No momento, as adições de funcionalidades restringem-se aos colaboradores do laboratório, os quais estão desenvolvendo um software de avaliação da concordância de registro inter- e intra-observadores, que trabalha com os arquivos gerados pelo módulo de Etografia do sistema [97].

Alguns aprimoramentos ainda podem ser realizados nas funções primárias do software. Por exemplo, o algoritmo de análise de atividade (subtração de imagem de plano de fundo) usado forneceu resultados

satisfatórios na segmentação do animal do plano de fundo. No entanto, alguns exemplos de comportamentos foram excluídos do conjunto de amostras, pois estavam prejudicados por mudanças intensas de iluminação, e/ou por artefatos adicionados ao vídeo em tempo de execução (e.g., fezes, urina). Nesses exemplos, o comprimento do animal e o ângulo de inclinação dessa reta apresentavam sucessivas falhas de cálculo, devido à presença de artefatos que desviavam a posição central da área do corpo do animal para uma posição fora de sua área. Apesar de prejudicarem os dados dos descritores, esses problemas da segmentação de imagens raramente têm influência sobre o rastreamento da atividade do animal, uma vez que o procedimento de rastreamento (*tracking*) reduz a área do corpo a apenas um ponto (partícula).

O estudo de descritores de imagens como método de pré-processamento para a identificação comportamental usando redes neurais começou com a avaliação de atributos estatísticos dos descritores de imagem (extraídos pelo ETHOWATCHER®), quanto a sua relevância na representação das diferenças existentes entre as categorias comportamentais. Usou-se da prova estatística não-paramétrica de Kruskal-Wallis para pré-selecionar os atributos relevantes. Conforme discutido, a avaliação da validade das escolhas dessa prova testou três conjuntos de atributos: atributos selecionados pela prova, apenas atributos de média e desvio padrão de cada descritor, e atributos não selecionados (considerados irrelevantes de acordo com a prova).

As redes neurais treinadas a partir dos atributos pré-selecionados e aquelas treinadas apenas com os atributos de média e desvio padrão dos descritores foram as que apresentaram os maiores desempenhos na identificação comportamental (Figuras 25-28). Além disso, na maioria dos casos não existia diferenças estatisticamente significantes entre os resultados desses dois grupos de classificadores (Mann-Whitney, $p > 0,05$). Verificou-se que nesses casos a pré-seleção também sugeria que os atributos de média e desvio padrão exibiam diferenças estatisticamente significantes entre categorias. Contudo, outros atributos que apresentavam diferenças mais significantes (menor nível de significância p) foram usados no conjunto de atributos selecionados.

Na comparação de desempenho entre os classificadores com atributos selecionados e aqueles usando os atributos não-selecionados, encontraram-se diferenças estatisticamente significantes entre os resultados, que sugerem um desempenho superior dos classificadores que usaram atributos selecionados. Esses resultados são confirmados

pela ausência de redes neurais treinadas em algumas dobras da validação cruzada do grupo não selecionado (níveis AUC de zero), além de desempenhos de identificação iguais ou inferiores a 50%, que retratam classificadores incapazes de diferenciar os comportamentos de interesse de outras categorias comportamentais (vide Figura 25).

A identificação comportamental usando os grupos de atributos relevantes apresentou um desempenho médio (AUC) superior a 94% na identificação de comportamentos no conjunto de validação (animais ingênuos a tratamentos, categorias: locomoção, imobilidade e exploração vertical). A identificação da limpeza corporal apresentou AUC média de 83,6% (vide Tabela 4). Esses resultados equivalem-se aos encontrados em trabalhos da área [17][21][26][92], superando-os em algumas categorias.

Em amostras de animais tratados com cafeína 2 mg/kg, obteve-se os seguintes níveis de desempenho (vide Tabela 5): $97,5 \pm 2,6\%$ (locomoção), $97,8 \pm 0,6\%$ (imobilidade), $71,8 \pm 3,0\%$ (exploração vertical), e $54,2 \pm 8,0\%$ (limpeza corporal). Apesar de não apresentarem diferenças estatisticamente significantes nos padrões de frequência e duração dos comportamentos e da análise de atividade (ao teste de Kruskal-Wallis), encontrou-se diferenças de desempenho estatisticamente significantes nos classificadores de exploração vertical e de limpeza corporal, na comparação com o desempenho na identificação desses comportamentos nas amostras de animais ingênuos ao tratamento (conjunto de validação). Embora não tenham sido vistas diferenças na frequência e duração dessas categorias, parece que essa dose de cafeína 2 mg/kg alterou os padrões morfológicos e cinemáticos das amostras.

Em segmentos de vídeo de experimentos com animais sob efeito da dose de 6 mg/kg, a locomoção foi identificada com AUC de $83,5 \pm 0,9\%$, a imobilidade com $85,2 \pm 1,8\%$, a exploração vertical com $67,0 \pm 2,0$, e a limpeza corporal com $78,0 \pm 10,0\%$ (Tabela 5). Alterações na frequência e na duração da exploração vertical, e na distância percorrida total sugerem que a cafeína alterou os padrões comportamentais desses animais (teste de Kruskal-Wallis, $p < 0,05$). Encontraram-se diferenças de desempenho estatisticamente significantes na comparação dos resultados desse cenário com os resultados no conjunto de validação, nos quatro comportamentos estudados, nos classificadores dos dois conjuntos de atributos testados (vide Tabela 5).

Essas diferenças estatísticas de desempenho podem ser consideradas como evidências de que a cafeína também alterou os padrões cinemáticos e morfológicos das categorias de comportamentos,

a ponto de em alguns casos torná-los ilegíveis aos classificadores treinados com amostras de animais ingênuos a tratamentos. Se considerada essa hipótese, a redução de desempenho encontrada na identificação da exploração vertical e da limpeza corporal na dose de 2 mg/kg também poderia ser efeito do tratamento. Desta forma, a ação dose-dependente desse tratamento poderia estar sendo encoberta, devido a baixas doses do tratamento não causarem efeitos estatisticamente significativos nos padrões de frequência e duração dos comportamentos, apenas em seus padrões cinemáticos e morfológicos.

Esse exemplo de aplicação das RNA assemelha-se ao emprego de máquinas de inteligência artificial na diferenciação dos padrões locomotores de animais afetados por doenças neuro-degenerativas em longo prazo, que eram imperceptíveis aos métodos de avaliação geralmente usados, até que o desenvolvimento da doença estivesse avançado [26].

Na avaliação de diferenças de desempenho entre os classificadores usando atributos selecionados e aqueles usando apenas a média e o desvio padrão na identificação de amostras comportamentais de animais tratados (cafeína, 2 mg/kg), aqueles usando atributos de média e desvio padrão apresentaram um desempenho superior na identificação da locomoção. Nos comportamentos de imobilidade e exploração vertical, o desempenho de identificação dos classificadores usando parâmetros selecionados foi marginalmente superior. Na limpeza corporal não houve diferenças estatisticamente significantes entre os dois grupos de atributos. Na dose de 6 mg/kg, a AUC média dos classificadores usando atributos selecionados foi superior nos comportamentos de exploração vertical e limpeza corporal, enquanto aqueles usando atributos de média e desvio padrão o foram para o comportamento de locomoção (Mann-Whitney, $p < 0,05$, vide Tabela 7). Esses resultados apóiam a validade da prova de Kruskal-Wallis como um método de seleção de parâmetros de entrada relevantes para a identificação comportamental usando RNA-MLP.

A avaliação de desempenho usando segmentos de 1 segundo de duração revelou que em algumas categorias de comportamento, esse período de tempo é insuficiente para resolver as ambiguidades existentes entre as categorias. Houve uma redução de desempenho estatisticamente significativa na comparação com o desempenho sobre o conjunto de validação em todos os comportamentos (Mann-Whitney, $p < 0,05$, vide Tabela 9). Os classificadores de locomoção apresentaram uma AUC de $88,80 \pm 0,02$ %, os de imobilidade $76,24 \pm 0,06$ %, os de exploração

vertical $68,96 \pm 0,03$ %, e os de limpeza corporal $51,60 \pm 0,09$ % (Tabela 8).

A análise detalhada dos resultados de cada classificador por amostra (Figura 29) revelou que ocorreram diagnósticos falso-positivos na identificação dos classificadores de limpeza corporal e exploração vertical, indicando que esse período de tempo tornou alguns eventos ambíguos aos classificadores. O classificador de exploração vertical marcou erroneamente algumas amostras de locomoção como de exploração, enquanto o classificador de limpeza corporal confundiu algumas amostras de imobilidade como de limpeza corporal.

Uma análise da duração das amostras usadas no conjunto de treinamento e avaliação demonstrou que as amostras de limpeza corporal e de exploração vertical tinham uma duração mínima superior ao período de tempo testado, enquanto as de locomoção e imobilidade possuíam (em alguns casos) uma duração mínima até inferior (vide histogramas das durações na Figura 24). Esses resultados evidenciaram que cada categoria tem um período de tempo mínimo distinto para sua diferenciação entre categorias semelhantes.

O desempenho dos testes discutidos foi medido através do índice AUC. No entanto, este trabalho comparou diferentes índices de desempenho na representação do desempenho de classificadores baseados em RNAs (vide Figura 30). Os classificadores usando atributos selecionados e aqueles usando apenas os atributos de média e desvio apresentaram tendências de resultado semelhantes entre os três índices, apesar de valores absolutos distintos (*e.g.*, o coeficiente Kappa subestima os resultados). No entanto, na identificação da limpeza corporal os classificadores usando atributos não selecionados apresentaram tendências de desempenho distintas. Enquanto a eficiência indicou um desempenho acima de 90% nas dobras da validação cruzada, o Kappa apresentou valores iguais e inferiores a 50%, e a AUC indicou uma dobra com desempenho superior a 80%, e oito delas com desempenho de zero.

Essas divergências se devem a dois fatores: o método de cálculo do índice de eficiência usado (eficiência absoluta), e a frequência de cada categoria comportamental no conjunto de exemplos deste trabalho. A eficiência absoluta é calculada sobre a frequência absoluta dos exemplos no conjunto de avaliação, sem ponderações dos acertos/erros em relação à quantidade de elementos de cada categoria. Desta forma, o índice retrata um desempenho tendencioso aos resultados das categorias com maior número de elementos. Como neste trabalho a limpeza corporal tinha a menor frequência de exemplos no conjunto de

treinamento e avaliação (vide Tabela 1), os erros de identificação da limpeza corporal tiveram pouca influência no valor final do índice, portanto recebendo um valor alto. Esses dados são confirmados pela baixa sensibilidade apresentada por esses classificadores, em oposição a uma alta especificidade (vide Figura 31).

Para o uso do índice eficiência na avaliação de desempenho recomenda-se a transformação dos índices VP, VN, FP, e FN em termos relativos às classes avaliadas (ponderação pelo número de elementos de cada classe). Uma alternativa a eficiência absoluta é o índice de eficiência balanceada [16], que não é influenciada pelas limitações discutidas. Escolheu-se os índices AUC e Kappa, pois esses índices ponderam seus valores de acordo com o desempenho de identificação do classificador em relação a ambas as classes avaliadas.

A Figura 30 ilustra classificadores de comportamentos em que os índices AUC e Kappa divergem na época de maior desempenho do classificador. Neste cenário de teste os classificadores selecionados usando o coeficiente Kappa apresentaram maior taxa de acerto absoluta que os da AUC. O índice AUC expressa o desempenho de registro (ou diagnóstico) de um evento em um conjunto de pontos de corte da função de ativação do neurônio de saída da rede neural. Em contrapartida, o Kappa e a eficiência quantificam o desempenho de apenas um ponto de corte de uma determinada época. Na maioria dos casos a AUC e o Kappa concordaram na escolha da época de maior desempenho de um classificador. Nesses casos o Kappa foi usado para determinar o ponto de corte de maior desempenho da época selecionada pela AUC.

A RNA MLP demonstrou-se uma escolha adequada para a identificação dos comportamentos estudados. Usou-se de uma estrutura reduzida, treinada a partir do algoritmo de RUMELHART *et al.* [57] [19], que obteve resultados semelhantes a métodos recentes (*e.g.*, máquinas de vetor de suporte), ditos de maior desempenho e menor tempo de treinamento. Crê-se que o desempenho alcançado deve-se ao uso de um procedimento integrado de treinamento e avaliação de desempenho, e de uma etapa de pré-seleção dos parâmetros de entrada.

O uso da validação cruzada para determinação das épocas de maior generalização evidenciou diferenças de desempenhos relativas a forma como o conjunto de exemplos era subdividido, principalmente nos casos em que os parâmetros de entrada não eram ótimos na representação das características dos eventos (conjunto de atributos não-selecionados, Figura 30). A pré-seleção dos parâmetros de entrada permitiu o projeto de classificadores com uma estrutura reduzida (três camadas), com (no máximo) oito neurônios na camada de entrada e sete

na camada oculta. O emprego de estruturas menores reduziu o tempo de treinamento, e permitiu deslocar o esforço empregado no treinamento para o teste de diferentes valores de parâmetros de treinamento (*e.g.*, taxa de aprendizagem, constante de momento).

Ao fim, não foram encontradas contribuições significativas diretamente relacionadas aos modelos de redes *Perceptron* de múltiplas camadas. Ressalta-se que o treinamento de redes distintas para cada categoria comportamental simplificou a avaliação da relevância dos parâmetros de entrada em cada categoria comportamental.

Deve-se considerar que embora o desempenho de identificação dos classificadores (treinados usando atributos estatisticamente relevantes) tenha decaído em alguns cenários de teste, ele ainda supera a concordância de registro de observadores humanos em muitos casos. Numa avaliação de concordância de registro entre observadores humanos iniciantes, o registro dos comportamentos de camundongos na arena de campo aberto alcançou um coeficiente Kappa máximo de 35,48% (e um índice de concordância pura de 76,71%, similar a eficiência absoluta) [97]. Apesar de não se referirem as mesmas amostras usadas para treinar as RNAS, esses valores demonstram a baixa confiabilidade dos resultados obtidos por observadores humanos pouco treinados.

6. CONCLUSÕES

O uso combinado de descritores de imagens e atributos estatísticos descreveu com sucesso as categorias comportamentais estudadas. Entre os descritores cinemáticos usados, destaca-se a variante usada do descritor de número de pixels alterados desenvolvida nesta tese (seção 3.3). A descrição da componente cinemática tem sido discutida recentemente como uma parte fundamental da representação e diferenciação de certas categorias comportamentos (*e.g.*, locomoção e imobilidade), que até então vinha sendo ignorada na representação do fenômeno comportamental [15].

Os desempenhos obtidos nos cenários de teste (conjunto de validação, amostras de animais tratados com cafeína, e segmentos de 1 segundo de amostras) permitem afirmar que o procedimento de pré-seleção adotado (Kruskal-Wallis e Mann-Whitney) é capaz de selecionar parâmetros relevantes na diferenciação das categorias comportamentais (de locomoção, imobilidade, exploração vertical, e limpeza corporal), válidos inclusive quando os padrões comportamentais estão alterados pelo uso de tratamento.

Salienta-se que apesar de ser determinante para diferenciar parâmetros relevantes de irrelevantes, o nível de significância (p) do teste estatístico usado não é recomendado para seleção entre os parâmetros relevantes. Sugere-se que nessa segunda etapa seja usado o desempenho do próprio classificador como forma de avaliar os parâmetros de entrada mais relevantes para o procedimento de identificação.

Evidencia-se o uso do desempenho de detecção de RNAs-MLP usando os descritores desenvolvidos, como parâmetro na avaliação de alterações em padrões de sinais comportamentais. Em uma sub-dose de tratamento (Cafeína 2 mg/kg) na qual os métodos usuais de análise comportamental não identificaram alterações (padrões de frequência, duração e latência), verificou-se uma redução de desempenho do classificador. Essas alterações foram confirmadas por ambos os métodos com o aumento da dose do tratamento.

Além da possibilidade de usar essas alterações de desempenho (geralmente reduções nos índices de acerto) como estimadores brutos da efetividade do tratamento em teste, esses resultados demonstram a necessidade de desenvolver algoritmos de inteligência computacional capazes de ignorar certo grau de alterações nos padrões em análise, como o faz o cérebro humano.

O treinamento de classificadores distintos para detecção de cada categoria comportamental demonstrou que cada categoria comportamental parece necessitar de uma quantidade de quadros (ou duração de tempo) mínima distinta para a identificação do comportamento.

Sobre o treinamento das RNAs, recomenda-se o índice AUC para seleção da época de maior desempenho da rede neural, e do coeficiente Kappa para seleção do ponto de corte de maior desempenho da função de ativação do neurônio de saída da rede. Na comparação de desempenho de classificadores, sugere-se o uso da eficiência balanceada aos índices de concordância e de eficiência absoluta. Essa escolha é importante quando as categorias estudadas têm frequências desbalanceadas no conjunto de exemplos usado para treinar ou avaliar o classificador. Essas diferenças irão direcionar o resultado do índice em relação ao desempenho nas categorias em maior quantidade no conjunto.

Em resumo, são vistas como contribuições desse trabalho: a descrição de eventos comportamentais (locomoção, imobilidade, exploração vertical, e limpeza corporal) através do conjunto de índices cinemáticos e morfológicos usados, a descrição do relacionamento desses descritores durante as amostras comportamentais usando atributos de estatística descritiva, a validação da prova de Kruskal-Wallis como procedimento de pré-seleção de parâmetros de entrada relevantes para diferenciação entre as categorias comportamentais, e a avaliação dos índices de desempenho Kappa, eficiência, e AUC na representação do desempenho de classificadores binários voltados a identificação comportamental. Por fim, o desenvolvimento da ferramenta ETHOWATCHER® para o estudo dos padrões comportamentais, além do registro e análise da atividade comportamental.

7. DESENVOLVIMENTOS FUTUROS

Os experimentos que forneceram os vídeos digitais usados para extrair os descritores foram obtidos de condições experimentais semelhantes (*e.g.*, níveis de iluminação, altura da câmera em relação à arena experimental). Em condições relativamente uniformes e semelhantes, usando animais de tamanho semelhante, esses classificadores devem ser capazes de reproduzir (ainda que parcialmente) os níveis de desempenho de identificação comportamental demonstrados. No entanto, se alterada a relação de tamanho entre a arena e o animal, ou as propriedades do dispositivo de aquisição (*e.g.*, taxa de aquisição de imagens), os classificadores terão seu desempenho reduzido, de forma similar aos testes usando segmento de um segundo dos comportamentos.

O estudo da quantidade de quadros necessária para identificar cada categoria comportamental deve ser estendido, verificando tanto a influência de janelas inferiores quanto superiores, como também diferentes taxas de aquisição, e variações de intensidade de cor de outros equipamentos de aquisição de imagem (*e.g.*, webcams, filmadoras). Houve casos em que ocorreram medições idênticas dos valores de descritores em quadros consecutivos. Essas repetições podem estar relacionadas tanto com a própria imobilidade do animal, como a deficiências na etapa de aquisição do vídeo (baixa taxa de aquisição), e artefatos do algoritmo de compactação.

Por fim, sugere-se a avaliação da identificação comportamental através de RNAs-MLP dinâmicas. Essas RNAs são capazes de mapear o relacionamento entre os descritores comportamentais extraídos de uma amostra, sem o uso de atributos de estatística descritiva. A comparação de resultado de ambas as abordagens permitirá o projeto de uma solução híbrida para o registro comportamental em ambiente experimental real.

APÊNDICE A – ETHOWATCHER®

Neste apêndice serão descritas as funções básicas da ferramenta ETHOWATCHER®. Todo o processo de configuração e execução da ferramenta é acompanhado por um tutorial iterativo, que baseado nas ações realizadas pelo usuário, indica quais os próximos passos a serem executados. A estrutura do tutorial inicia-se na tela inicial/configuração do sistema (Figura 32) e se estende pelas telas internas de cada módulo. É permitido ao usuário desabilitar o uso do tutor, pois à medida que o processo torna-se usual, tende a ser um incômodo para profissionais mais experientes.

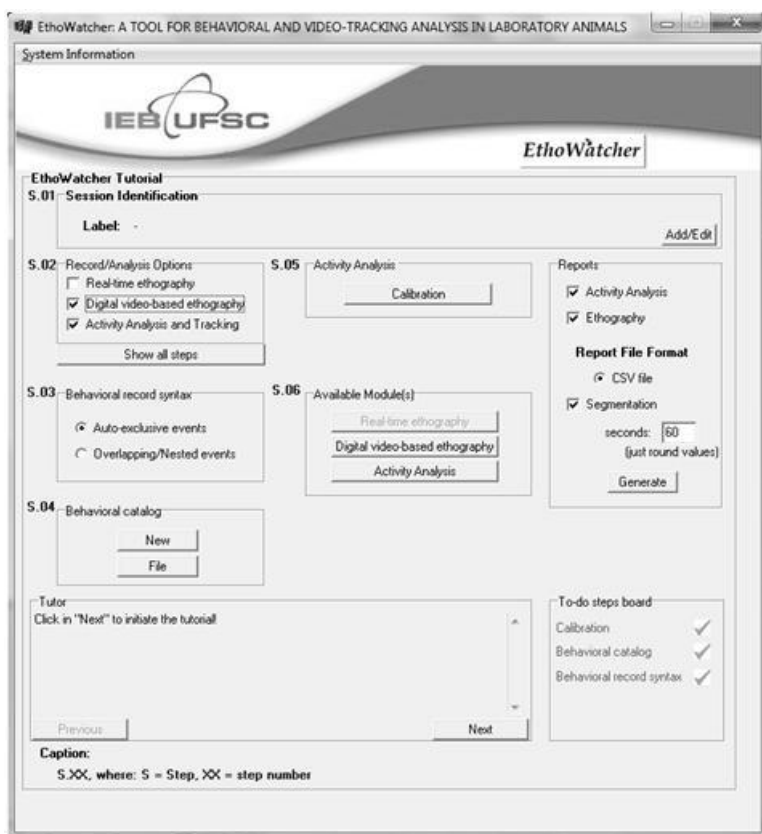


Figura 32. Tela Inicial – Tutorial

O projeto dos tutoriais seguiu os conceitos de orientação a objetos com uso das técnicas de herança e polimorfismo. Logo, procurou-se gerar um modelo de estruturas com os principais métodos de execução, de forma a manter uma nomenclatura padrão. As classes tutoras interagem com classes de controle, e desta forma, separam os métodos de interface com usuário das classes de negócio. A principal classe tutora é o Tutor da Ferramenta, responsável por coordenar todas as etapas de pré-processamento que forem necessárias com base na escolha do tipo de experimento que o usuário deseja realizar. O maior ciclo de etapas desse tutor contém 11 passos, que podem ser divididas em dois grupos, de configuração: identificação do experimento, sintaxe de registro, catálogo comportamental, calibração do instrumento; e de execução: Etografia Manual, Análise de Atividade e Etografia Automática.

As etapas de configuração e execução possuem tutores internos para auxiliar o usuário, com exceção da etapa de Sintaxe de Registro que está incluída no tutor principal. Um dos tutores mais importantes é o da Etapa de Calibração, pois frente aos outros, é o que a configuração é mais complexa e tem impacto direto na qualidade da análise de atividade. É subdividido em cinco etapas: seleção da imagem do animal, seleção do pixel de fundo, subtração do ambiente, erosão, e limitação da área de processamento, vide Figura 33.



Figura 33. Tela de Calibração do Instrumento

A etapa de seleção da imagem do animal corresponde à escolha da imagem que será base para calcular a segmentação do animal do fundo da imagem. A etapa de seleção do pixel de fundo permite a escolha da cor do pixel representativo do fundo da arena, e também é parte importante da diferenciação entre cores de fundo e a cor do animal. Em seguida, a etapa de subtração do ambiente demonstra o processo de segmentação em relação ao nível de subtração escolhido. A etapa posterior de erosão permite remover partes do animal que não foram retiradas na subtração, como bicos, caudas, etc. Por fim, a etapa de área de processamento permite delimitar a área de processamento a apenas uma parte da arena.

Enquanto a etapa de Calibração é essencial para o processamento da análise da atividade espacial do animal, a etapa de Criação de Catálogo o é para a Etografia. Nesta função, o experimentador define os sujeitos (animais) e as unidades comportamentais aceitas para o experimento que está configurando, o código alfa-numérico que representará cada unidade durante a Etografia, e a descrição contextual de comportamentos e sujeitos. A Figura 34 ilustra essa etapa.

System Information

IEB UFSC

EthoWatcher

New Behavioral Catalog

Back

Catalog Elements

Type: Category

Name: grooming

Code: g

Description: limpeza corporal auto-limpeza

Add Clean

Catalog at Design

Subject	Code
subject 01	s1
subject 02	s2

Category	Code
locomotion	l
inobility	i
rearing	r

New Catalog Save Catalog

Figura 34. Tela de Criação de Catálogo

Outra etapa de configuração, importante para o processo de Etografia, é a Sintaxe do Registro Etográfico. Esta configuração é realizada na Tela Inicial do sistema, Figura 32, e permite a escolha entre o registro de categorias auto-exclusivas ou sobrepostas. Registros auto-exclusivos impõem que ao entrar o código de uma categoria durante a Etografia, a categoria que estava em execução é automaticamente terminada. Na sintaxe de registros sobrepostas, várias categorias podem ocorrer ao mesmo tempo, ou seja, o início de uma categoria não implica no término de outra.

Existem duas etapas de execução na ferramenta ETHOWATCHER®, a Etografia e a Análise de atividade. Existem duas opções de Etografia, a do ambiente e a partir de vídeo digital. A Figura

35 ilustra tela de registro do comportamento a partir de um arquivo de vídeo. Um tutorial está presente nesta etapa para auxiliar o usuário a definir em que ponto do arquivo de vídeo será considerado o início da etapa de registro. Em muitos experimentos o início do registro etográfico não é necessariamente o início do vídeo. Logo, é permitido ao usuário definir esse momento. Com relação à sintaxe dos registros selecionada, quando for auto-exclusiva, o experimentador deve sinalizar o início de uma categoria inserindo o código da Categoria no campo Categoria, e automaticamente, o sistema finalizará a Categoria que estiver em andamento. Já quando a sintaxe for registros sobrepostos, o fim de um comportamento também deve ser sinalizado, sendo necessária a digitação do código de sua categoria novamente.

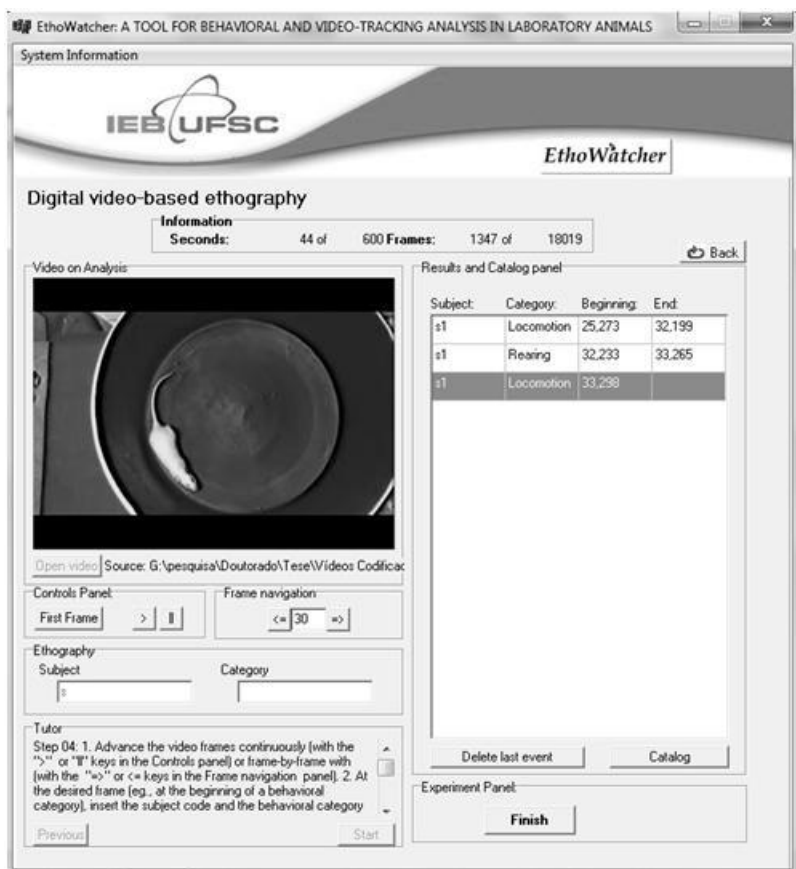


Figura 35. Tela de Etografia

A Figura 36 ilustra o relatório da função de auxílio na Etografia a partir do sistema, na sub-opção arquivo de vídeo.

	A	B	C	D	E	F	G	H
13	ETHOGRAPHY REPORT							
14	Analysis initiated at 0 (seconds) of the video file							
15	Analysis terminated at 34 (seconds) of the video file							
16								
17	Subject ID Test							
18	Time	Category	Duration(s)					
19	22,344	loc	0,999					
20	23,343	imo	1,765					
21	25,108	rear	0,733					
22	25,841	loc	0,599					
23	26,44	imo	0,733					
24	27,173	rear	0,4					
25	27,572	loc	0,699					
26	28,272	rear	0,766					
27	29,038	imo	0,533					
28	29,57	loc	2,431					
29	32,001	rear	2,264					
30								
31	RESULTS FOR THE ENTIRE PERIOD OF ANALYSIS							
32								
33	Category	Duration(: Freq.	Latency(s)					
34	loc	4,729	4	22,344				
35	imo	3,03	3	23,343				
36	rear	4,163	4	25,108				
37								

Figura 36. Exemplo de Relatório de Etografia

A análise de atividade de um experimento é realizada a partir da tela ilustrada pela Figura 37. Nesta etapa o usuário deve selecionar o vídeo a analisar, definir o ponto de início e de fim da análise, e o nome e local que será salva a figura contendo o traçado da análise de atividade. Ao fim da análise um botão com rótulo “Relatório” será exibido. Para visualizar o relatório é necessário salvá-lo primeiramente. O arquivo de relatório é salvo no formato aberto CSV (*Comma-separated values*), sendo compatível com a maioria das planilhas de cálculo do mercado, como EXCEL (MICROSOFT), e CALC (BROFFICE.ORG). A Figura 38 ilustra um exemplo de relatório de análise de atividade, com os valores de cada descritor comportamental em cada quadro do vídeo, e a Figura 39 ilustra dois gráficos produzidos na análise de atividade de animais. O primeiro gráfico (A) refere-se a um animal ingênuo a tratamento, e o segundo (B) a um animal tratado com cafeína. Vê-se

uma atividade aparentemente mais densa no gráfico segundo animal, que se explica pelo efeito estimulante causado pela cafeína.

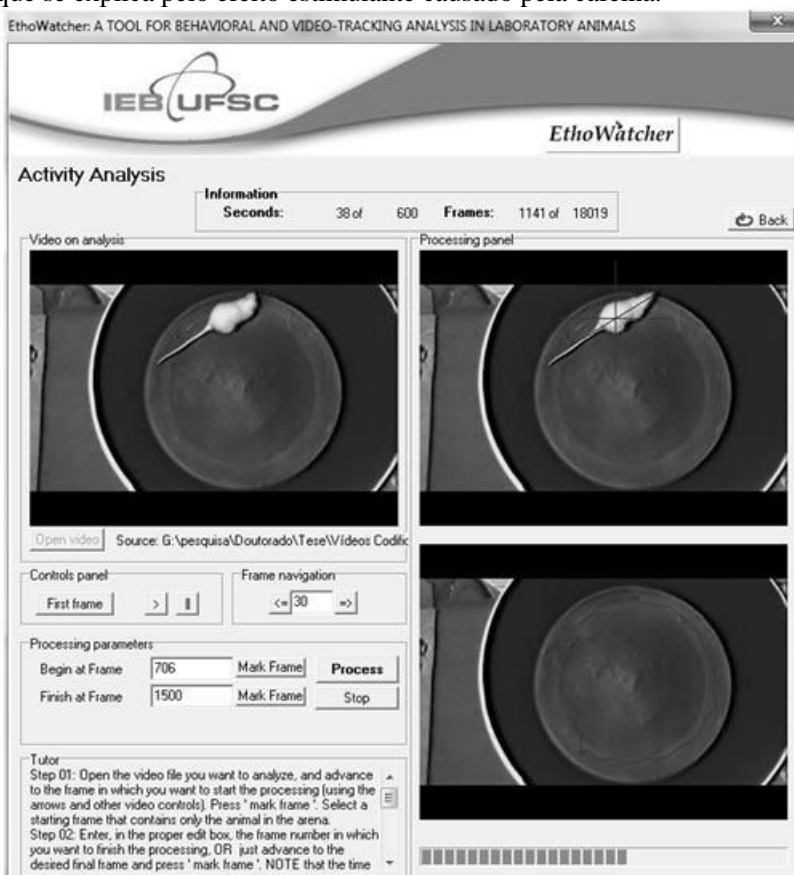


Figura 37. Tela de Análise de Atividade

report.csv - Microsoft Excel uso não comercial

Início Inserir Layout da Página Fórmulas Dados Revisão Exibição

Calibri 11 A A

Fonte Alinhamento Número Estilo

Formatação Condicional Inserir Excluir Formatar Formatar como Tabela Estilos de Célula Formatar Localizar e Filtrar Selecionar Edição

116

13 TRACKING - ACTIVITY REPORT

14

15 Vídeo fr C:\Users\Carlos\Desktop\Aluna Alexandre\Event\201006291448420042.avi

16 Start in 13 seconds 403 frames

17 End in 33 seconds 1000 frames

18 1 cm = 190 pixels

19 1 pixel = 0 cm2

20

Frames	Second	Number of Pixels of the animal's b of the animal	Area (in cm2)	Horizonta Vertical c	Distance travelled (in pixels)	Distance AA: Angle be Variation of the vertical i (in degrees) as compared to the former frame	Number of pixels with tone changes
21	403	13	17656	0.489	180	95	0
22	404	13	11826	0.328	178	114	21
23	405	13	1864	0.052	218	114	40
24	406	13	357	0.01	191	139	36
25	407	13	421	0.012	189	138	2
26	408	13	433	0.012	186	136	3
27	409	13	437	0.013	185	136	1
28	410	13	207	0.014	183	137	2
29	411	13	491	0.014	180	137	3
30	412	13	491	0.014	178	137	2
31	413	13	608	0.017	176	137	2
32	414	13	672	0.019	174	136	2
33	415	13	666	0.018	171	136	3
34	416	13	738	0.021	170	135	1
35	417	13	846	0.023	167	135	3
36	418	13	672	0.019	166	134	1
37	419	13	656	0.018	162	133	4
38	420	13	618	0.017	159	133	3
39	421	13	566	0.016	156	133	3
40	422	14	486	0.013	153	133	3
41	423	14	559	0.015	151	132	2
42	424	14	561	0.016	151	132	0
43	425	14	497	0.014	150	131	1

14 - 4 - 1 - report

Figura 38. Exemplo de Relatório de Atividade

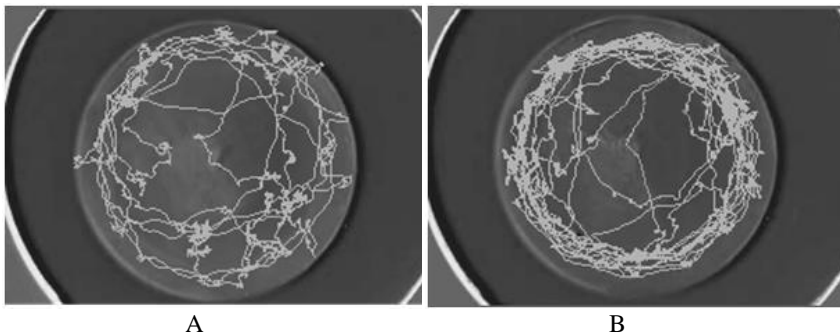


Figura 39. Imagens do rastreamento da atividade de dois ratos na arena de campo aberto. A imagem A ilustra o animal ingênuo ao tratamento, enquanto a imagem B ilustra o animal tratado com uma droga estimulante (caféina)

Foram descritas as telas consideradas de maior relevância para o uso do ETHOWATCHER® no registro e análise da atividade animal, sendo as outras suprimidas por lidarem com informações de controle do procedimento experimental, ou de uso do software.

APÊNDICE B – VALIDAÇÃO DOS DESCRITORES

O sistema ETHOWATCHER® e os descritores comportamentais apresentados na seção 3.3 foram avaliados em três situações distintas nas quais as medidas extraídas pelo sistema tinham dimensões ou comportamentos conhecidos. Essas situações foram: um objeto imóvel no centro e na periferia da arena de campo aberto (Teste 01); e um objeto quadrado e um objeto retangular realizando um movimento circular aproximadamente constante pela periferia da arena (Teste 02 e 03, respectivamente). Os arquivos de vídeos digitais de cada cenário de teste foram adquiridos através de uma filmadora SONY DCR-SR45, na resolução de 320x240 pixels, a uma taxa de aquisição de 30 quadros por segundo. A filmadora foi posicionada paralela ao centro da lâmpada de iluminação, sendo que ambas foram afixadas sobre a arena de campo aberto por um suporte.

O ambiente experimental foi uma arena do tipo campo aberto, em plástico, de formato circular, com altura de 45 cm, e diâmetro de 50,5 cm, pintada internamente com tinta acrílica preta. O ambiente estava iluminado diretamente por uma lâmpada fluorescente de 16 W, luz do dia (975 lumens), além da iluminação do ambiente. O iluminamento incidente na arena foi medido em alguns pontos, conforme Figura 40, nos quais apresentou os seguintes valores: 340 lux (1), 320 lux (3), 310 lux (4), 320 lux (5), e 274 lux (2). Essas medições de iluminamento foram realizadas usando o luxímetro MLM-1010 (MINIPA). Nas próximas seções serão descritos detalhadamente os testes de validação citados e os resultados obtidos. Os dados dos descritores estão expressos na forma de média \pm erro padrão da média (EPM).

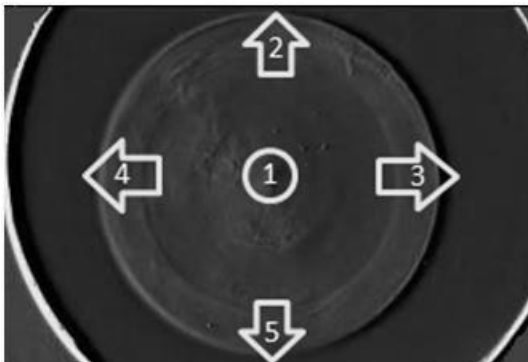


Figura 40. Referências dos pontos da arena de campo aberto nos quais foi medido o iluminamento

TESTE 01 – ESTABILIDADE DOS DESCRITORES EM OBJETOS PARADOS

Este teste adquiriu vídeos de um objeto posicionado em pontos onde os animais costumam transitar. Foram utilizadas duas caixas quadradas brancas de dimensões (comprimento x largura x altura, em centímetros) 6,1 x 6,1 x 6,0; e 7,2 x 7,2 X 5,1; posicionadas individualmente por 10 minutos no centro da arena e na periferia da arena. O posicionamento na periferia foi subdividido em dois períodos de igual duração (5 min.), em diferentes pontos da arena, para verificar diferenças relativas aos níveis de iluminamento. A aquisição de um período de 10 minutos de vídeo foi escolhida por ser uma duração comumente usada em experimentos com ratos e camundongos. A calibração do sistema foi realizada para cada um dos vídeos do experimento, sendo usada a caixa posicionada na periferia como referência (melhores resultados de segmentação).

Idealmente o sistema deveria estimar zero para os valores de descritores relacionados à cinemática, como: distância percorrida e número de pixels alterados; e manter constante os valores de descritores de orientação e tamanho, como o comprimento do animal, o número de pixels do animal e sua orientação em um plano cartesiano. No entanto, deformações métricas causadas pelas lentes da câmera na aquisição dos dados, ruídos inerentes ao processo de subtração de plano de fundo, e os métodos subsequentes de extração dos descritores, tendem a incluir variações nos dados.

Neste teste a posição do centróide manteve-se constante durante o experimento, e a distância percorrida foi estimada em zero. Apenas um caso apresentou um resultado diferente, porém ainda assim próximo a zero (caixa menor, centro), vide Tabela 11. O comprimento do animal e o descritor do ângulo dessa reta também mantiveram suas medidas constantes (variação angular zero), ainda que o valor absoluto apresentado diferisse para cada caixa entre os posicionamentos, vide Tabela 12 e Tabela 13.

Tabela 11. Distância Percorrida pelo objeto em relação aos Posicionamentos

Caixa \ Posição	Centro	Periferia 01	Periferia 02
Caixa 1 (menor)	0,003 ± 0,000432	0,000	0,000
Caixa 2 (maior)	0,000	0,000	0,000

Tabela 12. Comprimento do objeto (pixels) em relação aos Posicionamentos

Caixa \ Posição	Centro	Periferia 01	Periferia 02
Caixa 1 (menor)	64±0,000	59±0,000	64±0,000
Caixa 2 (maior)	77±0,000	73±0,000	65±0,000

Tabela 13. Ângulo de orientação da reta que caracteriza o comprimento do animal e sua variação entre quadros consecutivos em relação aos posicionamentos dos objetos

Caixa \ Posição	Centro	Periferia 1	Periferia 2
Caixa 1 (menor)	45(0,000)	41(0,000)	45(0,000)
Caixa 2 (maior)	47(0,000)	51(0,000)	39(0,000)

Entre todos os descritores, o número de pixels do animal apresentou a maior variação entre as diferentes medições de uma mesma caixa, sendo os diferentes posicionamentos de uma caixa estatisticamente diferentes (teste ANOVA de uma via). Apesar dessas diferenças, visualiza-se que na em dois de três casos há uma concordância (ou proximidade) de valores (*e.g.*, caixa menor nas regiões periféricas, e caixa maior no centro e na zona periférica dois). Essas variações tiveram pequena relevância quando analisados vídeos de ratos, no entanto elas devem ser estudadas mais detalhadamente para verificar seu efeito sobre descritores sobre animais menores (*e.g.*, camundongos).

Tabela 14. Número de Pixels do objeto em relação aos Posicionamentos

Caixa \ Posição	Centro	Periferia 1	Periferia 2
Caixa 1 (menor)	1053,06 ± 0,140	807,89 ± 0,016	795,72 ± 0,100
Caixa 2 (maior)	1345,86 ± 0,079	1260,47 ± 1,810	1324,58 ± 0,760

O descritor de número de pixels alterados aparentemente tão suscetível as influências do processo de segmentação quanto o número de pixels do animal teve melhores resultados para os propósitos da detecção da inatividade do objeto, apresentando uma média de menos de 1 pixels alterado por quadro no pior caso (zona periférica 1 - caixa maior), vide Tabelas 14 e 15.

Tabela 15. Número de Pixels alterados em relação aos Posicionamentos

Posição \ Caixa	Centro	Periferia 1	Periferia 2
Caixa 1 (menor)	0,289 ± 0,015	0,356 ± 0,021	0,336 ± 0,031
Caixa 2 (maior)	0,711 ± 0,016	0,939 ± 0,030	0,634 ± 0,012

Em resumo, os descritores cinemáticos atribuíram zero as suas medições, com pequenas variações, e os descritores de ângulo e comprimento mantiveram-se constantes em relação a cada posicionamento, apenas diferindo no valor absoluto dos dados. Foram observadas maiores variações nos dados extraídos pelos descritores para a caixa nº 2 (maior). No entanto, não foi possível comprovar estatisticamente essas diferenças. Considerou-se que para os propósitos deste trabalho os descritores comentados foram satisfatórios na representação de um objeto parado.

TESTE 02 – DISTÂNCIA PERCORRIDA REAL E ESPERADA

Este teste avaliou a distância percorrida estimada pelo ETHOWATCHER® para um objeto em movimento. Esse procedimento de validação baseou-se em um modelo anterior [43], que filmou o movimento realizado por um círculo branco pintado na borda de um disco de vinil em rotação. Usou-se um motor de corrente contínua, afixado no centro da arena de campo aberto para produzir o movimento de rotação. Sobre o eixo de rotação do motor foi afixada uma régua plástica transparente, de 30 cm de comprimento, a qual continha um quadrado branco na ponta oposta ao eixo (3 cm x 2,5 cm), conforme

Figura 41. A distância entre o ponto em que a régua estava afixada sobre o motor até o início do quadrado branco era de 17 cm, e a distância até o fim da régua (fim do quadrado branco) era de aproximadamente 20 cm. A régua foi coberta com papel preto fosco para evitar reflexões de iluminação. Para controlar o motor foi utilizada uma fonte de tensão regulável. Testaram-se três níveis de velocidades: 212,59 graus/segundo (0,74 m/s); 469,56 graus/segundo (1,67 m/s), e 739,73 graus/segundo (2,61 m/s). Uma marca foi adicionada ao fundo da arena para simbolizar o ponto de contagem do início e fim de cada volta, destacado por um círculo pontilhado da Figura 41.

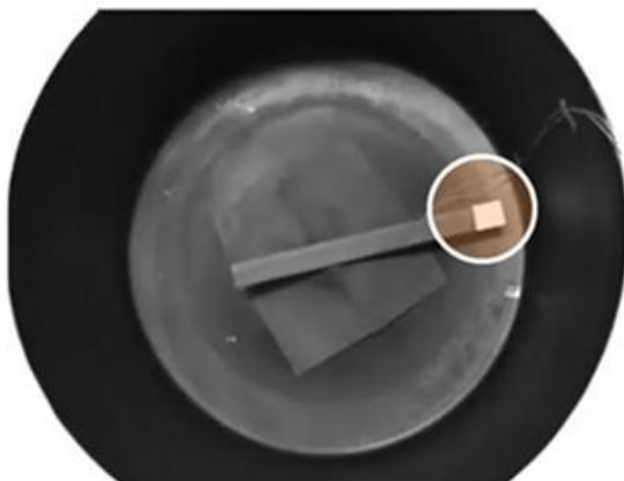


Figura 41. Ambiente de teste do Movimento Circular

Para cada velocidade testada foi adquirido um vídeo do objeto rastreado realizando cinco voltas, a contar da marcação do ponto de início. A Figura 42 ilustra o rastreamento dos vídeos do objeto em cada velocidade testada. Não houve falhas no rastreamento das três imagens, apenas variações da posição central do quadrado rastreado entre as velocidades de teste.

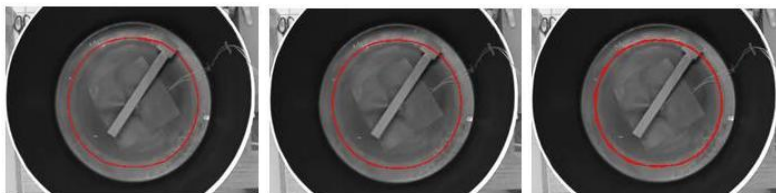


Figura 42. Rastreamento do objeto nas três velocidades testadas

A distância percorrida de referência foi calculada com base na distância entre o eixo do motor (centro da arena) até a borda do quadrado branco, localizada na ponta da régua. Como o sistema desenvolvido baseia-se na redução de todos os pixels que compõe o objeto a uma partícula (ponto médio), optou-se por determinar um intervalo de distâncias percorridas aceitáveis para um objeto que realizasse uma volta na circunferência. Considerando que o ponto médio do objeto seja estimado no início do quadro branco (ponto mais próximo do eixo do motor), a distância de referência seria em 106,81 cm (circunferência com 17 cm de raio). Caso o ponto médio fosse estimado próximo a ponta da régua, a distância de referência seria 128,80 cm (20,5 cm de raio).

Em resumo, a distância percorrida estimada deveria estar no intervalo de valores [106,81cm; 128,8 cm]. Para a distância média de uma volta na circunferência (raio de 18,7 cm; comprimento 117,49 cm), o objeto rastreado deve percorrer 587,46 cm em cinco voltas. Os gráficos da análise de atividade (Figura 42) e a distância percorrida estimada pelo software (634,53 cm em 69,4 cm/s; 639,33 cm em 153,3 cm/s e 636,55 cm em 241,7 cm/s) sugerem que mesmo velocidades acima das atingidas por um rato (de $2,92 \pm 0,31$ cm/s em animais ingênuos a tratamentos, e $4,05 \pm 0,36$ cm/s em animais tratados com cafeína), tiveram poucos efeitos sobre o algoritmo de rastreamento e as medições de distância. Além disso, não foram encontradas diferenças estatisticamente significantes entre as medições de diferentes velocidades (ANOVA de uma via, $p < 0,05$).

Além da distancia percorrida, o descritor de número de pixels alterados foi analisado neste teste como uma medida para representação da atividade do animal. Pressupôs-se que quanto mais o animal se locomovesse, um número maior de pixels seria alterado (maior atividade), e caso o animal (ou objeto) permanecesse parado, um número de pixels alterados tendendo a zero seria apresentado. Nos testes com objetos parados o número de pixels alterados teve uma média de menos de um pixel por quadro (caixas posicionadas em diferentes pontos da arena). No entanto, obtiveram-se valores que demonstram uma tendência de decrescimento nos valores do descritor conforme a velocidade aumenta ($276,355 \pm 0,67$ em 69,4 cm/s; $207,697 \pm 0,76$ em 153,36 cm/s; e $212,44 \pm 4,34$ em 241,69 cm/s).

Essa tendência de decrescimento pode ser explicada pela Figura 43, que ilustra a sobreposição das imagens resultantes do processo de segmentação do quadrado rastreado no período de tempo de uma volta.

Vê-se que conforme a velocidade de teste foi aumentada, a sobreposição temporal entre as áreas do objeto rastreado diminuiu. Esse comportamento inversamente proporcional do descritor de número de pixels alterados em relação à velocidade sugere que as variações de intensidade relativas a essa sobreposição influenciam de forma mais significativa o descritor que a variação de intensidade causada pelo aumento da velocidade.

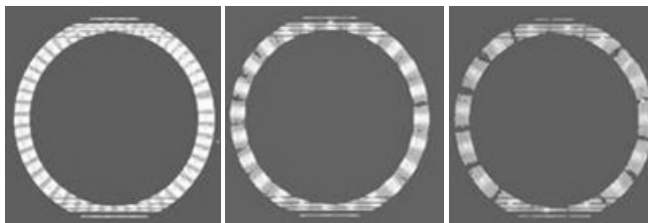


Figura 43. Sobreposição das imagens resultantes do procedimento de segmentação do objeto rastreado no período de uma volta, em cada uma das velocidades testadas (respectivamente: 69,4 cm/s; 153,36 cm/s; e 241,69 cm/s).

TESTE 03 – ORIENTAÇÃO DO OBJETO E VARIAÇÃO ANGULAR

Este teste usou dos mesmos parâmetros do Teste 02, apenas alterou-se o formato do objeto a ser rastreado, que passou de um quadrado para um retângulo. O objetivo deste teste foi verificar se o sistema era capaz de aferir a orientação de um objeto no plano cartesiano. O movimento circular realizado pelo objeto rastreado foi aferido em relação ao eixo das ordenadas, no intervalo de ângulos de -89 a 89 graus (vide procedimento de cálculo do descritor na seção 3.3). A projeção dos ângulos aferidos durante uma volta na circunferência gerou o sinal apresentado na Figura 44. Falhas na medição do ângulo do objeto são evidenciadas por perturbações na forma do sinal, e referem-se a problemas na medição da reta de maior comprimento do objeto, relativos a ruídos presentes na matriz resultante da segmentação do objeto do plano de fundo. O segundo gráfico da Figura 44 apresenta o sinal resultante das leituras dos ângulos referentes a 11,5 voltas à velocidade de 69,4 m/s. Ainda que as falhas existentes no desenho do sinal demonstrem que os procedimentos de cálculo de ângulo e comprimento possam ser aprimorados, o ruído existente não altera a morfologia do sinal de forma significativa. Vê-se que para cada volta na

circunferência o descritor de ângulo produz dois eventos (totalizando 23 eventos em 11,5 voltas).

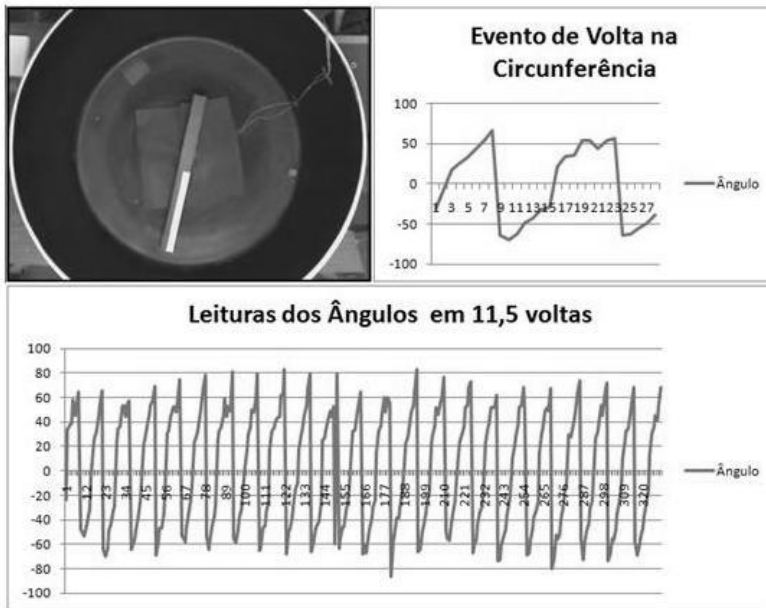


Figura 44. Procedimento de Teste de Movimento Circular com um Retângulo

CONCLUSÕES

Os testes descritos sugerem que os dados extraídos pelos descritores de imagem reproduzem os dados de referência, ainda que afetados pela presença de ruído. Nos testes com objetos imóveis a maior parte dos descritores demonstrou-se invariante em seus valores. Nos testes com objetos em movimento, os valores medidos estiveram dentro do intervalo de valores admitidos para o cenário (*e.g.*, distância percorrida entre 106,81cm e 128,8 cm), ou geraram eventos que caracterizavam o fenômeno medido (ângulo de orientação do comprimento do animal).

REFERÊNCIAS

- [01] LEVITIS, D. A.; LIDICKER, W. Z.; FREUND, G. Behavioural biologists do not agree on what constitutes behaviour. *Animal Behavior*, v. 78, 2009, p. 103-110.
- [02] MARTIN, P.; BATESON, P. *Measuring Behaviour*. New York: Cambridge University Press, 2007.
- [03] SIQUEIRA, L. O.; FERRARI, E. A. D. M.; MAIA, J. M. Sistema automático de análise comportamental em pombos. *Revista Brasileira de Engenharia Biomédica*, v. 22, n. 2, Agosto 2006, p. 93-105.
- [04] EISNER, T.; WILSON, E. O. *Animal Behavior: Readings from Scientific American*. San Francisco: W.H. Freeman and Co., 1975.
- [05] LEHNER, P. N. *Handbook of Ethological Methods*. 2^a Edição. ed. New York: Cambridge University Press, 1996.
- [06] LORENZ, K. Methods of approach to the problems of behavior. *The Harvey Lectures 1958-59*, New York, 1960, p. 60-103.
- [07] KELLY, A. E.; CADOR, M.; STINUS, L. Exploration and It's Measurement - A psychopharmacological perspective. *NeuroMethods*, v. 13, 1986, p. 95-144.
- [08] CUNHA, W. H. A. Alguns princípios de categorização, descrição e análise do comportamento. *Ciência e Cultura*, São Paulo, v. 28, n. 1, 1976, p. 15-24.
- [09] BEHAVIOR - Activity & Exploration. *TSE-Systems: Activity & Exploration*, 15/04 2011. Disponível em: <<http://www.tse-systems.com/products/behavior/activity/actimot2.htm>>. Acesso em: 02 Outubro 2010.
- [10] PHOTOBEM Activity System - Open Field. *San Diego Instruments*, 15/04/2011 2007. Disponível em: <http://sandiegoinstruments.com/prod_flexopen.htm>. Acesso em: 02 Junho 2010.
- [11] QUINN, L. P. *et al.* LABORAS: Initial pharmacological validation of a system allowing continuous monitoring of laboratory rodent behaviour. *Journal of Neuroscience Methods*, v. 130, n. 1, 2003, p. 83-92.
- [12] QUINN, L. P. *et al.* Further validation of LABORASTM using

- various dopaminergic manipulations in mice including MPTP-induced nigro-striatal degeneration. *Journal of Neuroscience Methods*, v. 156, 2006, p. 218–227.
- [13] SCHWARZ, S. *et al.* VIEWER: a program for visualising, recording, and. *Computer Methods and Programs in Biomedicine*, v. 67, 2002, p. 55–66.
- [14] SPRUIJT, B. M.; DEVISSER, L. Advanced behavioural screening: automated home cage ethology. *Drug Discovery Today: Technologies*, v. 3, n. 2, 2006, p. 231-237.
- [15] BENJAMINI, Y. *et al.* Ten ways to improve the quality of descriptions of whole-animal movement. *Neurosci. Biobehav. Rev.*, v. 34, n. 8, 2010, p. 1351-65.
- [16] FRÖLICH, H. *et al.* Automated classification of the behavior of rats in the forced swimming test with support vector machines. *Neural Networks*, v. 21, 2008, p. 92–101.
- [17] ROUSSEAU, J. B. I. *et al.* Classification of rat behavior with an image-processing method and a neural network. *Behavior Research Methods, Instruments & Computers*, v. 32, 2000, p. 63-71.
- [18] KULIKOV, A. V. *et al.* Automated analysis of antidepressants' effect in the forced swim test. *J Neurosci Methods*, v. 191, n. 1, 2010, p. 26-31.
- [19] HAYKIN, S. *Redes Neurais: Princípios e Práticas*. Tradução de Paulo Martins Engel. 2^a. ed. Porto Alegre: Bookman, 2001.
- [20] RODRIGUES, M. A. B. *Sistema para Detecção e Classificação Automática de Apnéias do Sono a partir de Registros Polissonográficos*. Florianópolis, 2002. Tese (Doutorado em Engenharia Elétrica) - Centro Tecnológico, Universidade de Santa Catarina.
- [21] HEEREN, D. J.; COOLS, A. R. Classifying postures of freely moving rodents with the help of Fourier descriptors and a neural network. *Behavior Research Methods, Instruments, & Computers*, v. 32, n. 1, 2000, p. 56-62.
- [22] BLUM, A.; LANGLEY, P. Selection of Relevant Features and examples in machine learning. *Artificial Intelligence*, v. 97, n. 1-2, 1997, p. 245-271.
- [23] SANCHEZ-MAROÑO, N.; ALONSO-BETANZOS, A.;

- CASTILLO, E. *A new wrapper method for feature subset selection*. Proc. of European Symposium on Artificial Neural Networks. Bruges: Belgium: [s.n.], 2005, p. 27-29.
- [24] DASH, M.; LIU, H. Feature Selection for Classification. *Intelligent Data Analysis*, v. 1, 1997, p. 131-156.
- [25] LEIGHTY, R. E. *et al.* Use of artificial neural networks to determine cognitive impairment and therapeutic effectiveness in Alzheimer's transgenic mice. *Journal of Neuroscience Methods*, v. 167, 2008, p. 358-366.
- [26] STEELE, A. D. *et al.* The power of automated high-resolution behavior analysis revealed by its application to mouse models of Huntington's and prion diseases, *Proc Natl Acad Sci U S A*. [S.l.]: [s.n.], 2007, p. 1983-8.
- [27] GONZALEZ, R. C.; WOODS, R. E. *Processamento de Imagens Digitais*. Tradução de Roberto Marcondes Cesar Junior e Luciano da Fontoura Costa. 1ª. ed. São Paulo: Blücher, 2000.
- [28] HALLIDAY, D.; RESNICK, R.; MERRILL, J. *Fundamentos de Física*. 3. ed. Sao Paulo: Moderna, v. 1, 1983.
- [29] PEDERIVA, C. N. *Etógrafo: um sistema de auxílio ao registro e à análise de dados comportamentais*. Florianópolis, 2005. Dissertação (Mestrado em Engenharia Elétrica) - Centro Tecnológico, Universidade de Santa Catarina.
- [30] PRUS, A. J.; JAMES, J. R.; ROSECRANS, J. A. Conditioned Place Preference. *In: BUCCAFUSCO, J. J. Methods of Behavior Analysis in Neuroscience*. Boca Raton: CRC Press, 2009. Cap. 4.
- [31] PELLOW, S. *et al.* Validation of open : closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *Journal of Neuroscience Methods*, v. 14, n. 3, Agosto 1985, p. 149-167.
- [32] PRUT, L.; BELZUNG, C. The openfield as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review. *European Journal of Pharmacology*, v. 463, 2003, p. 3-33.
- [33] WELKER, W. I. Escape, exploratory, and food-seeking responses of rats in a novel situation. *J. Comp. Physiol.*

- Psychol*, v. 52, 1959, p. 106-111.
- [34] KELLEY, A. E. *Methods in Behavioral Pharmacology*. [S.l.]: Elsevier Science Publishers, 1993.
- [35] UFJF. Laboratório de Neurofisiologia Cognitiva. *Laboratório de Neurofisiologia - UFJF*, 2009. Disponível em: <<http://www.neurofisiologia.ufjf.br/>>. Acesso em: 02 Junho 2009.
- [36] DA SILVA, E. S. *et al.* Behavioral and metabolic effects of central injections of orexins/hypocretins in pigeons (*Columba livia*). *Regulatory Peptides*, v. 147, 2008, p. 9-18.
- [37] HUTT, S. J.; HUTT, C. *Observação Direta e Medida do Comportamento*. Tradução de Carolina Martuscelli Bori. São Paulo: Editora da Universidade de São Paulo, 1974.
- [38] EIBL-EIBESFELDT, I. *Etología: Introducción al estudio comparado del comportamiento*. Tradução de Margarida Costa. Barcelona: Omega, 1979. 643 p.
- [39] OTTONI, E. B. EthoLog 2.2: A tool for the transcription and timing of behavior observation sessions. *Behavior Research Methods, Instruments & Computers*, v. 32, n. 3, 2000, p. 446-449.
- [40] PEDERIVA, C. N. *et al.* *Etógrafo: um sistema de auxílio ao registro e análise de dados comportamentais*. In: Anais da XX Reunião Anual da FESBE. Águas de Lindóia: [s.n.]. 2005.
- [41] BLUMSTEIN, D. T.; DANIEL, J. C. *Quantifying Behaviour the JWatcher Way*. Sunderland: Sinauer Associates, Inc. Publishers, 2007.
- [42] BOSE, R. C. *Proposta para um Sistema de Rastreamento e Análise de Movimentos em Animais de Laboratório*. Florianópolis, 2003. Dissertação (Mestrado em Engenharia Elétrica) - Centro Tecnológico, Universidade de Santa Catarina.
- [43] LIND, N. M. *et al.* Validation of a digital video tracking system for recording pig locomotor behaviour. *Journal of Neuroscience Methods*, v. 143, 2005, p. 123-132.
- [44] SPINK, A. J. *et al.* The EthoVision video tracking system: A tool for behavioral phenotyping of transgenic mice. *Physiology & Behavior*, v. 73, 2001, p. 731- 744.

- [45] WELCOME to ANY-maze. *ANY-maze*, 2009. Disponível em: <<http://www.anymaze.com/>>. Acesso em: 02 Junho 2009.
- [46] SMART Video Tracking System. *Bioseb In Vivo Research Instruments*, 2009. Disponível em: <http://www.bioseb.com/bioseb/anglais/default/item_id=63_cat_id=4_SMART%20Video%20Tracking%20System.php>. Acesso em: 02 Junho 2009.
- [47] VIEWER Features. *Biobserve*, 2009. Disponível em: <<http://www.biobserve.com/products/viewer/index.html>>. Acesso em: 02 Junho 2009.
- [48] ZURN, J. B.; JIANG, X.; MOTAI, Y. *Video-Based Rodent Activity Measurement Using Near-Infrared Illumination*. Proceedings of the IEEE Instrumentation and Measurement Technology Conference. Ottawa-Canada: [s.n.], 2005. p. 1928-1931.
- [49] METRIS, B. V. Laboras Product Information. *Metris B.V.*, 2008b. Disponível em: <http://www.metris.nl/en/products/laboras/laboras_information/>. Acesso em: 02 Junho 2009.
- [50] SILVA, J. A. *et al.* Modelos Ocultos de Markov Aplicados na Identificação de Comportamento de Camundongos. *Hifen*, Uruguaiana, v. 31, n. 59, 2007b.
- [51] YU, J.; TAN, J. Object density-based image segmentation and its applications in biomedical image analysis. *Computer methods and programs in biomedicine*, v.96, 2009, p. 193–204.
- [52] TWINING C.J.; TAYLOR, C.J.; COURTNEY P. Robust tracking and posture description for laboratory rodent using active shape models, *Behavior Research Methods, Instruments, & Computers*, v.33, n. 3, 2001, pp. 381-391.
- [53] JHUANG H. *et al.*, Automated home-cage behavioural phenotyping of mice. *Nature Communications*, 1:68, 2010.
- [54] DOLLAR, P.; RABAUD, V.; COTTRELL, G.; BELONGIE, S. *Behavior Recognition via Sparse Spatio-Temporal Features*. Proceedings of 2nd Joint IEEE International Workshop on VS-PETS, Beijing:IEEE, 2005, p. 15-16.
- [55] WANG Z.; YANG X.; XU Y.; YU S. CamShift guided particle filter for visual tracking. *Pattern Recognition Letters*, v. 30, n. 4, 2009.

- [56] PISTORI H. *et al.* Mice and larvae tracking using a particle filter with an auto-adjustable observation model. *Pattern Recognition Letters*, v. 31, 2010, pp. 337–346
- [57] RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Parallel Distributed Processing. In: _____ *Learning Internal Representations by Error Propagation*. Cambridge: M. I. T. Press, v. 1, 1986. p. 318-362.
- [58] BISHOP, C. M. *Neural Networks for Pattern Recognition*. New York: Oxford University Press, 1995.
- [59] KANDELL, A.; BUNKE, H. *Hybrid Methods in Pattern Recognition*. New Jersey: World Scientific Pub Co Inc, v. 47, 2002.
- [60] SOVIERZOSKI, M. A.; ARGOU, F. I. M.; AZEVEDO, F. M. D. Evaluation of ANN Classifiers During Supervised Training with ROC Analysis and Cross Validation. *Proceedings of the 2008 International Conference on BioMedical Engineering and Informatics*, v. 1, 2008, p. 274-278.
- [61] WIELAND, A.; LEIGHTON, R. Geometric analysis of neural network capabilities. *IEEE International Conference on Neural Networks*, San Diego, CA, v. III, 1987, p. 385-392.
- [62] AMARI, S.I. *et al.* Asymptotic Statistical Theory of Overtraining and Cross-Validation. *IEEE Transactions on Neural Networks*, v. 8, n. 5, Setembro 1997, p. 985.
- [63] BRICKLEY, M. R.; SHEPHERD, J. P. Performance of a Neural Network Trained to Make Third-molar Treatment-planning Decisions. *Neural Network in Dentistry Treatment Decisions*, v. 16, n. 2, Abril 1996, p. 153-160.
- [64] SOVIERZOSKI, M. A. *Avaliação de Descritores Morfológicos na Identificação de Eventos Epileptiformes*. Florianópolis, 2009. Tese (Doutorado em Engenharia Elétrica) - Centro Tecnológico, Universidade de Santa Catarina.
- [65] OBUCHOWSKI, N. A. Receiver Operating Characteristic Curves and Their Use in Radiology. *Radiology*, v. 229, n. 1, Outubro 2003, p. 3-8.
- [66] EBERHART, R. C.; DOBBINS, R. W. Neural network performance metrics for biomedical applications. *Computer-*

- Based Medical Systems. Proceedings of Third Annual IEEE Symposium*, 1990, p. 282–289.
- [67] PARK, S. H.; GOO, J. M.; JO, C.-H. Receiver Operationing Characteristic (ROC) Curve: Practical Review for Radiologists. *Korean J Radiol*, v. 5, Março 2004, p. 11-18.
- [68] SOVIERZOSKI, M. A.; ARGOUD, F. I. M.; AZEVEDO, F. M. Performance Evaluation of an ANN FF Classifier of Raw EEG Data using ROC Analysis. *Proceedings of the 2008 International Conference on BioMedical Engineering and Informatics*, v. 1, 2008, p. 332-336.
- [69] PENG, Y.; WU, Z.; JIANG, J. A novel feature seelction approach for biomedical data classification. *Journal of Biomedical Informatics*, v. 43, 2010, p. 15-23.
- [70] COHEN, J. Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* , v. XX, n. 1, 1960.
- [71] MARTINS, M. P.; GUIMARÃES, L. N. L.; FONSECA, L. M. G. *CLASSIFICADOR DE TEXTURAS POR REDES NEURAIS*. Anais do II Congresso Brasileiro de Computação – CBComp 2002. Itajaí: Univali, 2002, p.12.
- [72] CHEN, H. *et al.* Neural network-based computer-aided diagnosis in distinguishing malignant from benign solitary pulmonary nodules by computed tomography. *Chinese Medical Journal* , v. 120, n. 14, 2007, p. 1211-1215.
- [73] IFFAT, A. G.; LESLIE, S. S. Feature subset selection in large dimensionality domains. *Pattern Recognition*, 43, 2010, p. 5-13.
- [74] GHEYAS, I. A.; SMITH, L. S. Feature subset selection in large dimensionality domains. *Pattern Recognition*, v. 43, 2010, p. 5-13.
- [75] DASH, M.; LIU, H. Feature Selection for Classification. *Intelligent Data Analysis*, 1, 1997, p. 131-156.
- [76] XIN, J. *et al.* Machine Learning Techniques and Chi-Square Feature: Selection for Cancer Classification Using SAGE Gene Expression Profiles. *Lecture Notes in Computer Science*, 3916, 2006, p. 106–115.
- [77] LIAO, C.; LI, S.; LUO, Z. Gene selection using Wilcoxon rank sum test and support vector machine for cancer. *Lecture*

- Notes in Computer Science*, v. 4456, 2007, p. 57–66.
- [78] GREEN, T. A.; SCHENK, S. Dopaminergic Mechanism for Caffeine-Produced Cocaine Seeking in Rats. *Neuropsychopharmacology*, v. 26, n. 4, 2002, p. 422-430.
- [79] EL YACOUBI, M. *et al.* The stimulant effects of caffeine on locomotor behavior in mice are mediated through its blockade of adenosine A2A receptors. *British Journal of Pharmacology*, v. 129, 2000, p. 1465-1473.
- [80] SVENNINGSSON, P. *et al.* Antagonism of Adenosine A2A receptors underlies the behavioural activating effect of caffeine and is associated with reduced expression of messenger RNA for NGFI-A and NGFI-B in caudate putamen and nucleus accumbens. *Neuroscience*, v. 79, n. 3, 1997, p. 753-764.
- [81] LAU, C. E.; FALK, J. L. Dose-Dependent Surmountability of Locomotor Activity in Caffeine Tolerance. *Pharmacol. Biochem and Behav.*, v. 52, n. 1, 1995, p. 139-143.
- [82] LOCOMOTION. *EarthLife*, 2009. Disponível em: <<http://www.earthlife.net/mammals/locomotion.html>>. Acesso em: 02 Junho 2009.
- [83] GEYER, M. A. Approaches to the characterization of drug effects on locomotor activity in rodents. In: ADLER, M. W.; COWAN, A. *Testing and Evaluation of Drugs of Abuse*. New York: Wiley-Liss Inc, 1990. p. 81-99.
- [84] INTEL. "*Open Source Computer Vision Library - Reference Manual*. [S.l.]: Intel Corporation, 1999.
- [85] MACHADO, B. B. *et al.* Topolino: Software Livre para Automatização do Experimento de Campo Aberto. *XV Seminário de Computação - Seminco*, Blumenau, p. 20-22 Novembro 2006.
- [86] MEAGHER, R. K. Observer ratings: Validity and value as a tool for animal welfare research. *Applied Animal Behaviour Science*, v. 119, p. 1–14, 2009.
- [87] SILVA, J. A. *et al.* *Comparação de Descritores de Formas no Reconhecimento de Objetos*. Anais do III WVC - Workshop de Visão Computacional. São José do Rio Preto: [s.n.]. 2007a.
- [88] STATSOFT INC. (2011) Electronic Statistics Textbook.

WEB: <http://www.statsoft.com/textbook/>.

- [89] FREUND, J. E.; SIMON, A. G. *Estatística aplicada: economia, administração e contabilidade*. Tradução de Alfredo A Farias. 9. ed. Porto Alegre: Bookman, 2000.
- [90] KAZMIER, L. J. *Estatística aplicada à administração e economia*. Tradução de Adriano Silva Vale Cardoso. 4. edição. ed. Porto Alegre: Bookman, 2007, p. 392.
- [91] KOHAVI, R. *A study of cross-validation and bootstrap for accuracy estimation and model selection*. Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2. Montreal - Canada: International Joint Conference On Artificial Intelligence. 1995. p. 1137-1143.
- [92] GONZALEZ, L. P.; ARNALDO, C. M. Classification of drug-induced behaviors using a multi-layer feed-forward neural network. *Comput Methods Programs Biomed*, v. 40, n. 3, Jul 1993, p. 167-73.
- [93] SIEGEL, S. *Estatística não-paramétrica para as ciências do comportamento*. Tradução de Alfredo Alves Farias e Eva Nick. Sao Paulo: McGraw-Hill do Brasil, 1981.
- [94] SHAPIRO, S. S.; WILK, M. B. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, v. 52, Dezembro 1965, p. 591-611.
- [95] ROYSTON, P. Remark AS R94: A Remark on Algorithm AS 181: The W-test for Normality. *Journal of the Royal Statistical Society*, v. 44, 1995, p. 547-551. Series C (Applied Statistics).
- [96] LILLIEFOR, H. W. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association*, v. 62, n. 318, Junho 1967, p. 399-402.
- [97] GARCIA, V. A.; CRISPIM-JUNIOR, C. F.; MARINONETO, J. *Assessment of observers' stability and reliability - a tool for evaluation of intra- and inter-concordance in animal behavioral recordings*. Proceedings of 32nd Annual International Conference on the IEEE EMBS. Buenos Aires: IEEE, 2010, p. 6603-6606.