# Where is C3D Model used?

The C3D model is widely used in the field of computer vision for action recognition and video classification tasks (Martin, 2022). It has been applied in various domains, including sports gesture recognition (Martin, 2022), video event detection (Murtaza et al., 2019), depression detection from facial expressions (Melo et al., 2019), cross-domain action recognition (Shang et al., 2020), short video event detection (Li & Xu, 2020), and sleep monitoring (Torres et al., 2018).

In the domain of action recognition, the C3D model has been shown to outperform other models such as HOF and MBH in terms of motion-based descriptors (Wang & Schmid, 2013). It has also been used in combination with other models like ActionVLAD and P3D for improved performance (Shang et al., 2020). Additionally, the C3D model has been utilized in temporal activity detection Xu et al. (2020) and temporal action detection (Murtaza et al., 2019).

The C3D model has also found applications in other areas. For example, it has been used in the field of source separation for visually guided sound source separation (Zhu & Rahtu, 2020). In the medical field, C3D has been used for the detection of maternal antibodies and their clinical significance (Niewiesk, 2014). Furthermore, the C3D model has been applied in geological mapping and modeling, specifically in the context of creating a 2D and 3D geological map/model of Canada (Kemp, 2023).

Overall, the C3D model has proven to be a powerful tool for various tasks in computer vision, including action recognition, video classification, event detection, and source separation. Its effectiveness has been demonstrated in multiple domains, highlighting its versatility and wide range of applications.

References:

Kemp, E. (2023). Canada in 3d - national geological surveys committee update report.. https://doi.org/10.4095/331340

Li, P. and Xu, X. (2020). Recurrent compressed convolutional networks for short video event detection. Ieee Access, 8, 114162-114171. https://doi.org/10.1109/access.2020.3003939

Li, Q., Cheng, H., Zhou, Y., & Huo, G. (2016). Human action recognition using improved salient dense trajectories. Computational Intelligence and Neuroscience, 2016, 1-11. https://doi.org/10.1155/2016/6750459

Martin, P. (2022). 3d convolutional networks for action recognition: application to sport gesture recognition.. https://doi.org/10.48550/arxiv.2204.08460

Melo, W., Granger, E., & Hadid, A. (2019). Combining global and local convolutional 3d networks for detecting depression from facial expressions.. https://doi.org/10.1109/fg.2019.8756568

Murtaza, F., Yousaf, M., Velastin, S., & Yu, Q. (2019). End-to-end temporal action detection using bag of discriminant snippets. Ieee Signal Processing Letters, 26(2), 272-276. https://doi.org/10.1109/lsp.2018.2888758

Ng, J., Hausknecht, M., Vijayanarasimhan, S., Vinyals, O., Monga, R., & Toderici, G. (2015). Beyond short snippets: deep networks for video classification.. https://doi.org/10.48550/arxiv.1503.08909

Niewiesk, S. (2014). Maternal antibodies: clinical significance, mechanism of interference with immune responses, and possible vaccination strategies. Frontiers in Immunology, 5. https://doi.org/10.3389/fimmu.2014.00446

Shang, F., Han, T., Tian, F., Tao, J., & Gao, Z. (2020). A multimodal pairwise discrimination network for cross-domain action recognition. Ieee Access, 8, 143545-143557. https://doi.org/10.1109/access.2020.3014691

Torres, C., Fried, J., Rose, K., & Manjunath, B. (2018). A multiview multimodal system for monitoring patient sleep. Ieee Transactions on Multimedia, 20(11), 3057-3068. https://doi.org/10.1109/tmm.2018.2829162

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Wang, H. and Schmid, C. (2013). Action recognition with improved trajectories.. https://doi.org/10.1109/iccv.2013.441

Wang, H. and Schmid, C. (2013). Action recognition with improved trajectories.. https://doi.org/10.1109/iccv.2013.441

Wang, H., Oneață, D., Verbeek, J., & Schmid, C. (2015). A robust and efficient video representation for action recognition. International Journal of Computer Vision, 119(3), 219-238. https://doi.org/10.1007/s11263-015-0846-5

Xu, H., Yang, L., Sclaroff, S., Saenko, K., & Darrell, T. (2020). Spatio-temporal action detection with multi-object interaction.. https://doi.org/10.48550/arxiv.2004.00180

Zhu, L. and Rahtu, E. (2020). Visually guided sound source separation using cascaded opponent filter network.. <https://doi.org/10.48550/arxiv.2006.03028>

# how does HOF and MBH work with C3D model?

HOF (Histogram of Optical Flow) and MBH (Motion Boundary Histogram) are two motion-based descriptors that are commonly used in conjunction with the C3D model for action recognition tasks.

HOF captures the local motion information in a video by computing the optical flow between consecutive frames. It represents the motion information by dividing the image into spatial cells and quantizing the optical flow vectors into different bins. The resulting histogram represents the distribution of motion directions in the video Li et al. (2016). On the other hand, MBH focuses on capturing the motion boundaries in a video. It computes the gradient of the optical flow field and quantizes the gradient orientations into different bins. The resulting histogram represents the distribution of motion boundaries in the video (Wang & Schmid, 2013).

When used with the C3D model, HOF and MBH serve as complementary features to the spatial appearance information captured by the C3D model. The C3D model extracts spatiotemporal features from video frames using 3D convolutional layers (Tran et al., 2015). These features capture both appearance and motion information. By combining the appearance features from the C3D model with the motion-based features from HOF and MBH, a more comprehensive representation of the video is obtained, leading to improved action recognition performance (Ng et al., 2015).

The combination of HOF, MBH, and the C3D model allows for a more robust and efficient video representation for action recognition tasks. The C3D model captures the spatial and temporal information, while HOF and MBH capture the local motion and motion boundary information, respectively. This combination of features provides a more comprehensive representation of the video, enabling accurate action recognition (Wang et al., 2015).

Overall, HOF and MBH work in conjunction with the C3D model to capture motion information and enhance the performance of action recognition systems. By combining these motion-based descriptors with the spatial appearance features extracted by the C3D model, a more comprehensive and effective representation of videos is achieved.

References:

Kemp, E. (2023). Canada in 3d - national geological surveys committee update report.. https://doi.org/10.4095/331340

Li, P. and Xu, X. (2020). Recurrent compressed convolutional networks for short video event detection. Ieee Access, 8, 114162-114171. https://doi.org/10.1109/access.2020.3003939

Li, Q., Cheng, H., Zhou, Y., & Huo, G. (2016). Human action recognition using improved salient dense trajectories. Computational Intelligence and Neuroscience, 2016, 1-11. https://doi.org/10.1155/2016/6750459

Martin, P. (2022). 3d convolutional networks for action recognition: application to sport gesture recognition.. https://doi.org/10.48550/arxiv.2204.08460

Melo, W., Granger, E., & Hadid, A. (2019). Combining global and local convolutional 3d networks for detecting depression from facial expressions.. https://doi.org/10.1109/fg.2019.8756568

Murtaza, F., Yousaf, M., Velastin, S., & Yu, Q. (2019). End-to-end temporal action detection using bag of discriminant snippets. Ieee Signal Processing Letters, 26(2), 272-276. https://doi.org/10.1109/lsp.2018.2888758

Ng, J., Hausknecht, M., Vijayanarasimhan, S., Vinyals, O., Monga, R., & Toderici, G. (2015). Beyond short snippets: deep networks for video classification.. https://doi.org/10.48550/arxiv.1503.08909

Niewiesk, S. (2014). Maternal antibodies: clinical significance, mechanism of interference with immune responses, and possible vaccination strategies. Frontiers in Immunology, 5. https://doi.org/10.3389/fimmu.2014.00446

Shang, F., Han, T., Tian, F., Tao, J., & Gao, Z. (2020). A multimodal pairwise discrimination network for cross-domain action recognition. Ieee Access, 8, 143545-143557. https://doi.org/10.1109/access.2020.3014691

Torres, C., Fried, J., Rose, K., & Manjunath, B. (2018). A multiview multimodal system for monitoring patient sleep. Ieee Transactions on Multimedia, 20(11), 3057-3068. https://doi.org/10.1109/tmm.2018.2829162

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Wang, H. and Schmid, C. (2013). Action recognition with improved trajectories.. https://doi.org/10.1109/iccv.2013.441

Wang, H. and Schmid, C. (2013). Action recognition with improved trajectories.. https://doi.org/10.1109/iccv.2013.441

Wang, H., Oneață, D., Verbeek, J., & Schmid, C. (2015). A robust and efficient video representation for action recognition. International Journal of Computer Vision, 119(3), 219-238. https://doi.org/10.1007/s11263-015-0846-5

Xu, H., Yang, L., Sclaroff, S., Saenko, K., & Darrell, T. (2020). Spatio-temporal action detection with multi-object interaction.. https://doi.org/10.48550/arxiv.2004.00180

Zhu, L. and Rahtu, E. (2020). Visually guided sound source separation using cascaded opponent filter network.. https://doi.org/10.48550/arxiv.2006.03028

# Why is C3D model better than HOF or MBH models?

The C3D model is considered better than the HOF or MBH models for action recognition. The C3D model is a 3D Convolutional Neural Network (CNN) that is specifically designed for spatiotemporal feature learning (Tran et al., 2015). It has been shown to outperform state-of-the-art methods on multiple benchmarks and is comparable to the best methods on other benchmarks (Tran et al., 2015). The C3D model is trained on a large dataset of videos, including 1 million examples from the Sports-1M dataset (Carreira & Zisserman, 2017). This large training dataset allows the C3D model to learn rich spatiotemporal features that are effective for action recognition (Carreira & Zisserman, 2017).

In contrast, the HOF and MBH models focus on capturing local motion information (Wang & Schmid, 2013). While they can achieve similar performance to the C3D model when the foreground optical flow is correct (Wang & Schmid, 2013), they may not be as effective in capturing the overall spatiotemporal patterns in videos. The HOF and MBH models also rely on handcrafted features such as Histograms of Oriented Gradients (HOG) and optical flow (Xu & Wong, 2017). These handcrafted features may not be as expressive or discriminative as the learned features in the C3D model.

Furthermore, the C3D model has been shown to capture long-range temporal structures, which are important for understanding complex actions (Qiu et al., 2017). This ability to model temporal dependencies is crucial for accurate action recognition. In contrast, the HOF and MBH models may not capture long-range temporal structures as effectively.

Overall, the C3D model is considered better than the HOF or MBH models for action recognition due to its ability to learn rich spatiotemporal features, capture long-range temporal structures, and outperform state-of-the-art methods on multiple benchmarks.

References:

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Qiu, Z., Yao, T., & Mei, T. (2017). Learning spatio-temporal representation with pseudo-3d residual networks.. https://doi.org/10.1109/iccv.2017.590

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Wang, H. and Schmid, C. (2013). Action recognition with improved trajectories.. https://doi.org/10.1109/iccv.2013.441

Xu, T. and Wong, E. (2017). Learning temporal structures for human activity recognition.. <https://doi.org/10.5244/c.31.160>

# What are the real life usages of C3D model?

The C3D model has various real-life applications in different domains. In the field of medical research, the acquisition of C3d-binding activity by de novo donor-specific HLA antibodies has been correlated with graft loss in nonsensitized pediatric kidney recipients (Comoli et al., 2016). This suggests that the C3D model can be used to identify and predict the risk of graft rejection in kidney transplantation.

In immunology, the structure-guided identification of C3d residues essential for its binding to complement receptor 2 (CD21) has been studied (Clemenza & Isenman, 2000). This research highlights the importance of understanding the molecular interactions between C3d and CD21, which can be facilitated by the C3D model. Such knowledge can contribute to the development of therapeutic interventions for immune-related disorders.

In the field of computer science, a variant of the C3D model called MV-C3D has been proposed for spatially correlated multi-view 3D convolutional neural networks (Xuan et al., 2019). This model demonstrates its adaptability in real-world scenarios, particularly in the analysis of 3D-rotated real image datasets. The MV-C3D model can be utilized in computer vision applications, such as object recognition and scene understanding.

Additionally, the molecular adjuvant C3d has been used to enhance antibodies to the human immunodeficiency virus type 1 envelope (Green et al., 2003). This research suggests that the C3D model can be employed in the development of vaccine strategies and immunotherapies for viral infections.

Overall, the C3D model finds applications in diverse fields, including medical research, immunology, computer science, and virology. Its ability to analyze spatiotemporal features makes it a valuable tool for understanding complex phenomena and improving various processes in these domains.

References:

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Clemenza, L. and Isenman, D. (2000). Structure-guided identification of c3d residues essential for its binding to complement receptor 2 (cd21). The Journal of Immunology, 165(7), 3839-3848. https://doi.org/10.4049/jimmunol.165.7.3839

Comoli, P., Cioni, M., Tagliamacco, A., Quartuccio, G., Innocente, A., Fontana, I., … & Ginevri, F. (2016). Acquisition of c3d‐binding activity by de novo donor‐specific hla antibodies correlates with graft loss in nonsensitized pediatric kidney recipients. American Journal of Transplantation, 16(7), 2106-2116. https://doi.org/10.1111/ajt.13700

Green, T., Montefiori, D., & Ross, T. (2003). Enhancement of antibodies to the human immunodeficiency virus type 1 envelope by using the molecular adjuvant c3d. Journal of Virology, 77(3), 2046-2055. https://doi.org/10.1128/jvi.77.3.2046-2055.2003

Qiu, Z., Yao, T., & Mei, T. (2017). Learning spatio-temporal representation with pseudo-3d residual networks.. https://doi.org/10.1109/iccv.2017.590

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Wang, H. and Schmid, C. (2013). Action recognition with improved trajectories.. https://doi.org/10.1109/iccv.2013.441

Xu, T. and Wong, E. (2017). Learning temporal structures for human activity recognition.. https://doi.org/10.5244/c.31.160

Xuan, Q., Li, F., Liu, Y., & Xiang, Y. (2019). Mv-c3d: a spatial correlated multi-view 3d convolutional neural networks. Ieee Access, 7, 92528-92538. <https://doi.org/10.1109/access.2019.2923022>

# Where is P3D model used?

The P3D model is a versatile tool that is used in various scientific disciplines to model and simulate complex systems. In biology, the P3D model is used in genome-wide association studies (GWAS) (Zhang et al., 2010). It is also used in the modeling of Li-ion batteries to understand their dynamic operation and heat generation (Diddens et al., 2022). In geology, the P3D model is used for hydraulic fracturing simulations and to evaluate the best intervals for hydraulic fracturing (Amiri et al., 2021; Nguyen et al., 2021). It is also used in the modeling of solid electrolyte interfaces in Li-ion batteries (Carelli et al., 2019). In petroleum engineering, the P3D model is widely used in hydraulic fracturing simulations (Chen et al., 2021). In chemistry, the P3D model is used for 3D reconstructions and structure determination of small molecules (Vollnhals & Wirtz, 2018; Opakua & Zweckstetter, 2021). It is also used in the modeling and design of semi-solid flow batteries (Chayambuka et al., 2019). In genetics, the P3D model is used in association mapping and phenotypic diversity studies (Gunjača et al., 2021; Rueda et al., 2022).

References:

Amiri, H., Ramezanzadeh, A., Chamanzad, M., & Parhizgar, M. (2021). Recognizing the best intervals for hydraulic fracturing using a new fracability index. Journal of Petroleum Exploration and Production Technology, 11(8), 3193-3201. https://doi.org/10.1007/s13202-021-01173-7

Carelli, S., Quarti, M., Yagci, M., & Bessler, W. (2019). Modeling and experimental validation of a high-power lithium-ion pouch cell with lco/nca blend cathode. Journal of the Electrochemical Society, 166(13), A2990-A3003. https://doi.org/10.1149/2.0301913jes

Chayambuka, K., Fransaer, J., & Dominguez-Benetton, X. (2019). Modeling and design of semi-solid flow batteries. Journal of Power Sources, 434, 226740. https://doi.org/10.1016/j.jpowsour.2019.226740

Chen, B., Dias, R., Sun, Y., Bai, J., Thomas, H., Dutko, M., … & Li, C. (2021). A review of hydraulic fracturing simulation. Archives of Computational Methods in Engineering, 29(4), 1-58. https://doi.org/10.1007/s11831-021-09653-z

Diddens, D., Appiah, W., Mabrouk, Y., Heuer, A., Vegge, T., & Bhowmik, A. (2022). Modeling the solid electrolyte interphase: machine learning as a game changer?. Advanced Materials Interfaces, 9(8). https://doi.org/10.1002/admi.202101734

Gunjača, J., Carović-Stanko, K., Lazarević, B., Vidak, M., Petek, M., Liber, Z., … & Šatović, Z. (2021). Genome-wide association studies of mineral content in common bean. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.636484

Nguyen, H., Lee, J., & Elraies, K. (2021). Review of pseudo-three-dimensional modeling approaches in hydraulic fracturing. Journal of Petroleum Exploration and Production Technology, 12(4), 1095-1107. https://doi.org/10.1007/s13202-021-01373-1

Opakua, A. and Zweckstetter, M. (2021). Extending the applicability of p3d for structure determination of small molecules. Magnetic Resonance, 2(1), 105-116. https://doi.org/10.5194/mr-2-105-2021

Rueda, D., Awika, H., Bedre, R., Kandel, D., Mandadi, K., Crosby, K., … & Avila, C. (2022). Phenotypic diversity and association mapping of ascorbic acid content in spinach. Frontiers in Genetics, 12. https://doi.org/10.3389/fgene.2021.752313

Vollnhals, F. and Wirtz, T. (2018). Correlative microscopy in 3d: helium ion microscopy-based photogrammetric topography reconstruction combined with in situ secondary ion mass spectrometry. Analytical Chemistry, 90(20), 11989-11995. https://doi.org/10.1021/acs.analchem.8b02530

Zhang, Z., Ersoz, E., Lai, C., Todhunter, R., Tiwari, H., Gore, M., … & Buckler, E. (2010). Mixed linear model approach adapted for genome-wide association studies. Nature Genetics, 42(4), 355-360. <https://doi.org/10.1038/ng.546>

# P3D model and Li-ion batteries

The P3D model has been applied in the context of Li-ion batteries in various studies. One study by Chayambuka et al. (2019) focused on the modeling and design of semi-solid flow batteries (SSFBs) using a pseudo three-dimensional (P3D) model. The P3D model was used to simulate species diffusion as the only transport mechanism in the P3D active particles of the SSFBs. The model allowed for the illustration of time-dependent voltage profiles, current distribution, and state-of-charge distribution in the active particles.

Another study by Siczek et al. (2020) compared Li-S and Li-ion batteries in cooperation with a photovoltaic array. The authors utilized models based on electrical equivalent circuits for both types of batteries during their calculations. This study did not mention the use of the P3D model.

In a study by (Son & Du, 2019), a two-dimensional deterministic thermal dynamic model was employed to describe a cylindrical Li-ion battery cell. This study did not explicitly mention the use of the P3D model.

Additionally, Hughes et al. (2020) investigated the strain and chemistry evolution of the solid electrolyte-electrode interface in Li-ion batteries using a model system consisting of LiPON-LiCoO2. This study did not specifically mention the P3D model.

Overall, the P3D model has been utilized in the modeling and design of semi-solid flow batteries, but its direct application in other Li-ion battery studies is not explicitly mentioned in the available references.

References:

Amiri, H., Ramezanzadeh, A., Chamanzad, M., & Parhizgar, M. (2021). Recognizing the best intervals for hydraulic fracturing using a new fracability index. Journal of Petroleum Exploration and Production Technology, 11(8), 3193-3201. https://doi.org/10.1007/s13202-021-01173-7

Carelli, S., Quarti, M., Yagci, M., & Bessler, W. (2019). Modeling and experimental validation of a high-power lithium-ion pouch cell with lco/nca blend cathode. Journal of the Electrochemical Society, 166(13), A2990-A3003. https://doi.org/10.1149/2.0301913jes

Chayambuka, K., Fransaer, J., & Dominguez-Benetton, X. (2019). Modeling and design of semi-solid flow batteries. Journal of Power Sources, 434, 226740. https://doi.org/10.1016/j.jpowsour.2019.226740

Chayambuka, K., Fransaer, J., & Dominguez-Benetton, X. (2019). Modeling and design of semi-solid flow batteries. Journal of Power Sources, 434, 226740. https://doi.org/10.1016/j.jpowsour.2019.226740

Chen, B., Dias, R., Sun, Y., Bai, J., Thomas, H., Dutko, M., … & Li, C. (2021). A review of hydraulic fracturing simulation. Archives of Computational Methods in Engineering, 29(4), 1-58. https://doi.org/10.1007/s11831-021-09653-z

Diddens, D., Appiah, W., Mabrouk, Y., Heuer, A., Vegge, T., & Bhowmik, A. (2022). Modeling the solid electrolyte interphase: machine learning as a game changer?. Advanced Materials Interfaces, 9(8). https://doi.org/10.1002/admi.202101734

Gunjača, J., Carović-Stanko, K., Lazarević, B., Vidak, M., Petek, M., Liber, Z., … & Šatović, Z. (2021). Genome-wide association studies of mineral content in common bean. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.636484

Hughes, L., Savitzky, B., Gopal, C., Herring, P., Anapolsky, A., Ophus, C., … & Minor, A. (2020). Investigating strain and chemistry evolution of the solid electrolyte-electrode interface via 4d-stem and eels. Microscopy and Microanalysis, 26(S2), 150-151. https://doi.org/10.1017/s1431927620013598

Nguyen, H., Lee, J., & Elraies, K. (2021). Review of pseudo-three-dimensional modeling approaches in hydraulic fracturing. Journal of Petroleum Exploration and Production Technology, 12(4), 1095-1107. https://doi.org/10.1007/s13202-021-01373-1

Opakua, A. and Zweckstetter, M. (2021). Extending the applicability of p3d for structure determination of small molecules. Magnetic Resonance, 2(1), 105-116. https://doi.org/10.5194/mr-2-105-2021

Rueda, D., Awika, H., Bedre, R., Kandel, D., Mandadi, K., Crosby, K., … & Avila, C. (2022). Phenotypic diversity and association mapping of ascorbic acid content in spinach. Frontiers in Genetics, 12. https://doi.org/10.3389/fgene.2021.752313

Siczek, K., Siczek, K., Piersa, P., Adrian, Ł., Szufa, S., Obraniak, A., … & Bogusławski, G. (2020). The comparative study on the li-s and li-ion batteries cooperating with the photovoltaic array. Energies, 13(19), 5109. https://doi.org/10.3390/en13195109

Son, J. and Du, Y. (2019). Model-based stochastic fault detection and diagnosis of lithium-ion batteries. Processes, 7(1), 38. https://doi.org/10.3390/pr7010038

Vollnhals, F. and Wirtz, T. (2018). Correlative microscopy in 3d: helium ion microscopy-based photogrammetric topography reconstruction combined with in situ secondary ion mass spectrometry. Analytical Chemistry, 90(20), 11989-11995. https://doi.org/10.1021/acs.analchem.8b02530

Zhang, Z., Ersoz, E., Lai, C., Todhunter, R., Tiwari, H., Gore, M., … & Buckler, E. (2010). Mixed linear model approach adapted for genome-wide association studies. Nature Genetics, 42(4), 355-360. <https://doi.org/10.1038/ng.546>

# P3D model in genome studies

The use of three-dimensional (3D) models in genome studies has emerged as a promising approach to understanding the organization and function of the genome in health and disease (Babu & Fullwood, 2015). Genome geometry has been shown to play a crucial role in various cellular contexts that impact human health (Babu & Fullwood, 2015). In cancer translational medicine, for example, 3D genome organization has provided new opportunities for understanding the mechanisms underlying cancer development and progression (Babu & Fullwood, 2015).

One of the latest approaches in genome-wide association studies (GWAS) is the use of population parameters previously determined (P3D) (Li & Zhu, 2013). P3D is a two-step approach that first estimates population parameters, such as genetic and residual variance, using a compressed mixed linear model (CMLM) without testing SNP effects (Li & Zhu, 2013). In the second step, P3D estimates the testing SNP effects and random genetic effects using the population parameters determined in the first step (Li & Zhu, 2013). This approach improves computational efficiency by eliminating the need for multiple iterations of the mixed model equation (Gonzales & Palmer, 2014).

P3D has been applied in various GWAS studies, including those focused on non-model crops like Sesamum indicum (Berhe et al., 2021). In these studies, P3D has been used in combination with other improved single-locus genome scans MLM-based approaches, such as efficient mixed model association (EMMA) and random-SNP-effect mixed linear model (MRMLM) (Berhe et al., 2021). These approaches have demonstrated their effectiveness in identifying significant markers associated with traits of interest (Rueda et al., 2022).

The P3D approach has also been extended to other fields, such as hydraulic fracturing and semi-solid flow batteries (Chayambuka et al., 2019; Nguyen et al., 2021). In hydraulic fracture modeling, the pseudo-three-dimensional (P3D) approach has been used as a compromise between two-dimensional and planar three-dimensional models (Nguyen et al., 2021). In the modeling of semi-solid flow batteries, a P3D model has been developed to illustrate time-dependent voltage profiles, current distribution, and state-of-charge distribution in the active particles (Chayambuka et al., 2019).

In conclusion, the use of P3D models in genome studies, particularly in GWAS, has shown promise in improving computational efficiency and identifying significant markers associated with traits of interest. The P3D approach eliminates the need for multiple iterations of the mixed model equation, making it a valuable tool in various fields, including cancer translational medicine, non-model crop studies, hydraulic fracturing, and semi-solid flow batteries.

References:

Babu, D. and Fullwood, M. (2015). 3d genome organization in health and disease: emerging opportunities in cancer translational medicine. Nucleus, 6(5), 382-393. https://doi.org/10.1080/19491034.2015.1106676

Berhe, M., Dossa, K., You, J., Mboup, P., Diallo, I., Diouf, D., … & Wang, L. (2021). Genome-wide association study and its applications in the non-model crop sesamum indicum. BMC Plant Biology, 21(1). https://doi.org/10.1186/s12870-021-03046-x

Chayambuka, K., Fransaer, J., & Dominguez-Benetton, X. (2019). Modeling and design of semi-solid flow batteries. Journal of Power Sources, 434, 226740. https://doi.org/10.1016/j.jpowsour.2019.226740

Gonzales, N. and Palmer, A. (2014). Fine-mapping qtls in advanced intercross lines and other outbred populations. Mammalian Genome, 25(7-8), 271-292. https://doi.org/10.1007/s00335-014-9523-1

Li, G. and Zhu, H. (2013). Genetic studies: the linear mixed models in genome-wide association studies. The Open Bioinformatics Journal, 7(1), 27-33. https://doi.org/10.2174/1875036201307010027

Nguyen, H., Lee, J., & Elraies, K. (2021). Review of pseudo-three-dimensional modeling approaches in hydraulic fracturing. Journal of Petroleum Exploration and Production Technology, 12(4), 1095-1107. https://doi.org/10.1007/s13202-021-01373-1

Rueda, D., Awika, H., Bedre, R., Kandel, D., Mandadi, K., Crosby, K., … & Avila, C. (2022). Phenotypic diversity and association mapping of ascorbic acid content in spinach. Frontiers in Genetics, 12. <https://doi.org/10.3389/fgene.2021.752313>

# I3D Model usage

The usage of I3D models, or Inflated 3D Convolutional Neural Networks, has gained significant attention in various fields, including computer science, medicine, and population genomics (Carreira & Zisserman, 2017; Jeong et al., 2022; Begun et al., 2007). I3D models are an extension of state-of-the-art image classification architectures, where filters and pooling kernels are inflated into 3D, enabling spatiotemporal analysis (Carreira & Zisserman, 2017).

In the field of computer science, I3D models have been applied to action recognition in videos. introduced the Two-Stream Inflated 3D ConvNets, which demonstrated the effectiveness of I3D models in capturing spatiotemporal features for action recognition tasks (Carreira & Zisserman, 2017). The inflation of filters and pooling kernels into 3D allows for deep and naturally spatiotemporal classifiers (Carreira & Zisserman, 2017).

In the medical field, I3D models have been utilized for temporal analysis of videofluoroscopic swallowing studies. compared the performance of ResNet3D variants with that of VGG and I3D models in analyzing swallowing patterns (Jeong et al., 2022). The I3D model showed promising results in accurately analyzing temporal changes in swallowing dynamics (Jeong et al., 2022).

Furthermore, I3D models have also been employed in population genomics studies. conducted a population genetic analysis of Drosophila simulans using whole-genome shotgun sequencing (Begun et al., 2007). By comparing the genomic data of D. simulans with closely related species, such as D. melanogaster and D. yakuba, the study provided insights into polymorphism and divergence within the Drosophila genus (Begun et al., 2007).

In conclusion, I3D models have proven to be valuable tools in various domains. In computer science, they have been successful in action recognition tasks, capturing spatiotemporal features in videos. In the medical field, I3D models have shown promise in analyzing temporal changes in swallowing patterns. Additionally, in population genomics studies, I3D models have contributed to understanding genetic variation and divergence within species.

References:

Babu, D. and Fullwood, M. (2015). 3d genome organization in health and disease: emerging opportunities in cancer translational medicine. Nucleus, 6(5), 382-393. https://doi.org/10.1080/19491034.2015.1106676

Begun, D., Holloway, A., Stevens, K., Hillier, L., Poh, Y., Hahn, M., … & Langley, C. (2007). Population genomics: whole-genome analysis of polymorphism and divergence in drosophila simulans. Plos Biology, 5(11), e310. https://doi.org/10.1371/journal.pbio.0050310

Berhe, M., Dossa, K., You, J., Mboup, P., Diallo, I., Diouf, D., … & Wang, L. (2021). Genome-wide association study and its applications in the non-model crop sesamum indicum. BMC Plant Biology, 21(1). https://doi.org/10.1186/s12870-021-03046-x

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Chayambuka, K., Fransaer, J., & Dominguez-Benetton, X. (2019). Modeling and design of semi-solid flow batteries. Journal of Power Sources, 434, 226740. https://doi.org/10.1016/j.jpowsour.2019.226740

Gonzales, N. and Palmer, A. (2014). Fine-mapping qtls in advanced intercross lines and other outbred populations. Mammalian Genome, 25(7-8), 271-292. https://doi.org/10.1007/s00335-014-9523-1

Jeong, S., Kim, J., Park, J., Baek, S., & Yang, S. (2022). Application of deep learning technology for temporal analysis of videofluoroscopic swallowing studies.. https://doi.org/10.21203/rs.3.rs-2311543/v1

Li, G. and Zhu, H. (2013). Genetic studies: the linear mixed models in genome-wide association studies. The Open Bioinformatics Journal, 7(1), 27-33. https://doi.org/10.2174/1875036201307010027

Nguyen, H., Lee, J., & Elraies, K. (2021). Review of pseudo-three-dimensional modeling approaches in hydraulic fracturing. Journal of Petroleum Exploration and Production Technology, 12(4), 1095-1107. https://doi.org/10.1007/s13202-021-01373-1

Rueda, D., Awika, H., Bedre, R., Kandel, D., Mandadi, K., Crosby, K., … & Avila, C. (2022). Phenotypic diversity and association mapping of ascorbic acid content in spinach. Frontiers in Genetics, 12. <https://doi.org/10.3389/fgene.2021.752313>

# what is Two-Stream Inflated 3D ConvNets?

Two-Stream Inflated 3D ConvNets (I3D) is a model that builds upon state-of-the-art image classification architectures by inflating their filters and pooling kernels into 3D, resulting in deep spatiotemporal classifiers (Carreira & Zisserman, 2017). The I3D model is based on a two-stream network, which uses image-based methods to represent videos from RGB and optical flow streams (Hu et al., 2022). It is designed to improve the accuracy of action recognition tasks by capturing both appearance and motion information.

The concept of inflating 2D ConvNets into 3D ConvNets is a key component of the I3D model. This inflation process allows for pre-training with ImageNet models, which helps to leverage the knowledge learned from large-scale image classification tasks (Wang et al., 2018). By inflating the filters and pooling kernels into 3D, the I3D model is able to capture spatiotemporal information in videos, making it more suitable for action recognition compared to 2D ConvNets (Tran et al., 2015).

The I3D model has been applied in various domains, including violence detection (Freire-Obregón et al., 2021), driver behavior analysis (Chen et al., 2020), and action recognition (Yang et al., 2022). In the context of action recognition, the I3D model has been shown to achieve high accuracies when pre-trained on the Kinetics dataset (Hara et al., 2018). To further improve the accuracy of the I3D model, techniques such as sparse regularization Yang et al. (2022) and nonmaximum suppression Tsai et al. (2020) have been proposed.

The I3D model has also been extended and modified in different ways. For example, the SRI3D model proposes a novel two-stream inflated 3D ConvNet based on sparse regularization for action recognition (Yang et al., 2022). The Cross-Enhancement Transform Two-Stream 3D ConvNets algorithm considers the action distribution characteristics on a specific dataset to enhance the performance of the I3D model (Cao et al., 2019). Additionally, efficient parallel inflated 3D convolution architectures have been developed to improve the computational efficiency of the I3D model (Huang et al., 2020).

In summary, Two-Stream Inflated 3D ConvNets (I3D) is a model that inflates the filters and pooling kernels of state-of-the-art image classification architectures into 3D, allowing for the capture of spatiotemporal information in videos. The I3D model has been applied in various domains and has been extended and modified to improve its performance and efficiency. It has shown promising results in action recognition tasks, especially when pre-trained on large-scale datasets like Kinetics.

References:

Cao, D., Xu, L., & Zhang, D. (2019). Cross-enhancement transform two-stream 3d convnets for action recognition.. https://doi.org/10.1145/3371425.3371491

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Chen, J., Lee, C., Huang, P., & Lin, C. (2020). Driver behavior analysis via two-stream deep convolutional neural network. Applied Sciences, 10(6), 1908. https://doi.org/10.3390/app10061908

Freire-Obregón, D., Barra, P., Castrillón-Santana, M., & Marsico, M. (2021). Inflated 3d convnet context analysis for violence detection. Machine Vision and Applications, 33(1). https://doi.org/10.1007/s00138-021-01264-9

Hara, K., Kataoka, H., & Satoh, Y. (2018). Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet?.. https://doi.org/10.1109/cvpr.2018.00685

Hu, W., Liu, H., Du, Y., Yuan, C., Li, B., & Maybank, S. (2022). Interaction-aware spatio-temporal pyramid attention networks for action classification. Ieee Transactions on Pattern Analysis and Machine Intelligence, 44(10), 7010-7028. https://doi.org/10.1109/tpami.2021.3100277

Huang, Y., Guo, Y., & Gao, C. (2020). Efficient parallel inflated 3d convolution architecture for action recognition. Ieee Access, 8, 45753-45765. https://doi.org/10.1109/access.2020.2978223

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Tsai, J., Hsu, C., Wang, W., & Huang, S. (2020). Deep learning-based real-time multiple-person action recognition system. Sensors, 20(17), 4758. https://doi.org/10.3390/s20174758

Wang, L., Li, W., & Gool, L. (2018). Appearance-and-relation networks for video classification.. https://doi.org/10.1109/cvpr.2018.00155

Yang, Z., An, G., Zhang, R., Zheng, Z., & Ruan, Q. (2022). Sri3d: two‐stream inflated 3d convnet based on sparse regularization for action recognition. Iet Image Processing, 17(5), 1438-1448. <https://doi.org/10.1049/ipr2.12725>

# I3D model real life usage?

The I3D model has found real-life usage in various domains, particularly in action recognition tasks. One example is its application in surgical workflow recognition. utilized the I3D model, specifically the Inflated 3D ConvNet architecture, to recognize surgical workflows in Sleeve Gastrectomy procedures (Zhang et al., 2021). They trained the model using focal loss and achieved accurate recognition of different surgical steps.

Another real-life usage of the I3D model is in violence detection. applied the I3D model, which is based on a two-stream network, for context analysis in violence detection. By inflating the filters and pooling kernels into 3D, the I3D model was able to capture spatiotemporal information and achieve accurate violence detection.

The I3D model has also been employed in driver behavior analysis. used a modified version of the I3D model, where two-dimensional ConvNets were used instead of three-dimensional ConvNets, to analyze driver behavior. The spatial and temporal ConvNet streams were pre-trained using the large-scale ImageNet dataset, demonstrating the versatility of the I3D model in different applications.

Furthermore, the I3D model has been widely used in general action recognition tasks. introduced the I3D model and demonstrated its effectiveness in action recognition by achieving state-of-the-art results on the Kinetics dataset (Carreira & Zisserman, 2017). The I3D model has been applied in various action recognition scenarios, such as sports analysis, human activity recognition, and gesture recognition.

In summary, the I3D model has been successfully applied in real-life scenarios such as surgical workflow recognition, violence detection, driver behavior analysis, and general action recognition tasks. Its ability to capture spatiotemporal information by inflating 2D ConvNets into 3D has proven to be effective in achieving high accuracy in these applications.

References:

Cao, D., Xu, L., & Zhang, D. (2019). Cross-enhancement transform two-stream 3d convnets for action recognition.. https://doi.org/10.1145/3371425.3371491

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Chen, J., Lee, C., Huang, P., & Lin, C. (2020). Driver behavior analysis via two-stream deep convolutional neural network. Applied Sciences, 10(6), 1908. https://doi.org/10.3390/app10061908

Freire-Obregón, D., Barra, P., Castrillón-Santana, M., & Marsico, M. (2021). Inflated 3d convnet context analysis for violence detection. Machine Vision and Applications, 33(1). https://doi.org/10.1007/s00138-021-01264-9

Hara, K., Kataoka, H., & Satoh, Y. (2018). Can spatiotemporal 3d cnns retrace the history of 2d cnns and imagenet?.. https://doi.org/10.1109/cvpr.2018.00685

Hu, W., Liu, H., Du, Y., Yuan, C., Li, B., & Maybank, S. (2022). Interaction-aware spatio-temporal pyramid attention networks for action classification. Ieee Transactions on Pattern Analysis and Machine Intelligence, 44(10), 7010-7028. https://doi.org/10.1109/tpami.2021.3100277

Huang, Y., Guo, Y., & Gao, C. (2020). Efficient parallel inflated 3d convolution architecture for action recognition. Ieee Access, 8, 45753-45765. https://doi.org/10.1109/access.2020.2978223

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Tsai, J., Hsu, C., Wang, W., & Huang, S. (2020). Deep learning-based real-time multiple-person action recognition system. Sensors, 20(17), 4758. https://doi.org/10.3390/s20174758

Wang, L., Li, W., & Gool, L. (2018). Appearance-and-relation networks for video classification.. https://doi.org/10.1109/cvpr.2018.00155

Yang, Z., An, G., Zhang, R., Zheng, Z., & Ruan, Q. (2022). Sri3d: two‐stream inflated 3d convnet based on sparse regularization for action recognition. Iet Image Processing, 17(5), 1438-1448. https://doi.org/10.1049/ipr2.12725

Zhang, B., Ghanem, A., Simes, A., Choi, H., & Yoo, A. (2021). Surgical workflow recognition with 3dcnn for sleeve gastrectomy. International Journal of Computer Assisted Radiology and Surgery, 16(11), 2029-2036. <https://doi.org/10.1007/s11548-021-02473-3>

# What is the difference between Inflated 3D ConvNet and I3D model?

The Inflated 3D ConvNet (I3D) model is a spatiotemporal classifier that builds upon state-of-the-art image classification architectures by inflating their filters and pooling kernels into 3D (Carreira & Zisserman, 2017). This inflation process allows the model to capture both spatial and temporal information in videos, making it suitable for action recognition tasks (Carreira & Zisserman, 2017). The I3D model is based on a two-stream network, which consists of two sub-networks: one for appearance and one for motion (Freire-Obregón et al., 2021). The appearance sub-network processes RGB frames, while the motion sub-network processes optical flow frames (Freire-Obregón et al., 2021). By combining the information from both streams, the I3D model can effectively recognize actions in videos (Freire-Obregón et al., 2021).

On the other hand, the term "Inflated 3D ConvNet" refers to a general concept of inflating 2D ConvNets into 3D ConvNets (Tran et al., 2015). This concept was first introduced by in their work on spatiotemporal feature learning with 3D Convolutional Networks (Tran et al., 2015). They found that 3D ConvNets are more suitable for spatiotemporal feature learning compared to 2D ConvNets (Tran et al., 2015). They also proposed a specific architecture called C3D (Convolutional 3D), which consists of homogeneous layers with small 3x3x3 convolution kernels Tran et al. (2015). The C3D architecture, with a simple linear classifier, outperformed state-of-the-art methods on multiple benchmarks (Tran et al., 2015).

In summary, the I3D model is a specific implementation of the Inflated 3D ConvNet concept (Carreira & Zisserman, 2017). It inflates the filters and pooling kernels of state-of-the-art image classification architectures into 3D, allowing it to capture both spatial and temporal information for action recognition (Carreira & Zisserman, 2017). The I3D model is based on a two-stream network and has been widely used in various applications, such as violence detection, surgical workflow recognition, and activity recognition (Freire-Obregón et al., 2021; Zhang et al., 2021; Zhang et al., 2020). On the other hand, the term "Inflated 3D ConvNet" refers to the general concept of inflating 2D ConvNets into 3D ConvNets, which was first introduced by (Tran et al., 2015). The C3D architecture proposed by is one example of an Inflated 3D ConvNet Tran et al. (2015).

References:

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Freire-Obregón, D., Barra, P., Castrillón-Santana, M., & Marsico, M. (2021). Inflated 3d convnet context analysis for violence detection. Machine Vision and Applications, 33(1). https://doi.org/10.1007/s00138-021-01264-9

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Zhang, B., Ghanem, A., Simes, A., Choi, H., & Yoo, A. (2021). Surgical workflow recognition with 3dcnn for sleeve gastrectomy. International Journal of Computer Assisted Radiology and Surgery, 16(11), 2029-2036. https://doi.org/10.1007/s11548-021-02473-3

Zhang, Y., Gu, Y., Marsic, I., Zheng, Y., & Burd, R. (2020). Video-based concurrent activity recognition for trauma resuscitation.. <https://doi.org/10.1109/ichi48887.2020.9374399>

# What is Focal Loss?

The Focal Loss is a specialized loss function that addresses the problem of class imbalance in object detection tasks (Lin et al., 2017). In object detection, the majority of the image consists of background regions, while the regions containing objects of interest are relatively sparse. This class imbalance can lead to the model being biased towards the background class and performing poorly on detecting the foreground objects (Lin et al., 2017).

The Focal Loss introduces a modulating term to the standard cross-entropy loss function, which down-weights the loss assigned to well-classified examples (Lin et al., 2017). This modulating term focuses the training on hard examples, which are typically the minority class samples (Lin et al., 2017). By down-weighting the loss for easy examples, the Focal Loss helps the model to pay more attention to the challenging examples and improve its performance on the minority class (Lin et al., 2017).

The formulation of the Focal Loss is a generalized form that can simultaneously address the imbalance between foreground and background classes, as well as the imbalance among different foreground classes (Lin et al., 2017). This makes it particularly effective for dense object detection tasks, where there can be multiple foreground classes with varying levels of imbalance (Lin et al., 2017).

The Focal Loss has been successfully applied in various computer vision tasks, such as object detection and logo positioning (Lin et al., 2017; Wang, 2023). It has also been used in combination with ensemble learning algorithms to improve credit risk prediction in unbalanced datasets (Chen et al., 2023). The effectiveness of the Focal Loss lies in its ability to handle class imbalance and focus the model's attention on hard examples, leading to improved performance in challenging scenarios.

References:

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Chen, H., Yang, C., Du, M., & Zhang, Y. (2023). Research on credit risk prediction under unbalanced dataset based on ensemble learning. Mathematical Problems in Engineering, 2023, 1-18. https://doi.org/10.1155/2023/2927393

Freire-Obregón, D., Barra, P., Castrillón-Santana, M., & Marsico, M. (2021). Inflated 3d convnet context analysis for violence detection. Machine Vision and Applications, 33(1). https://doi.org/10.1007/s00138-021-01264-9

Lin, T., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection.. https://doi.org/10.1109/iccv.2017.324

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3d convolutional networks.. https://doi.org/10.1109/iccv.2015.510

Wang, Y. (2023). Research on vehicle logo positioning method based on deep learning.. https://doi.org/10.1117/12.2684296

Zhang, B., Ghanem, A., Simes, A., Choi, H., & Yoo, A. (2021). Surgical workflow recognition with 3dcnn for sleeve gastrectomy. International Journal of Computer Assisted Radiology and Surgery, 16(11), 2029-2036. https://doi.org/10.1007/s11548-021-02473-3

Zhang, Y., Gu, Y., Marsic, I., Zheng, Y., & Burd, R. (2020). Video-based concurrent activity recognition for trauma resuscitation.. <https://doi.org/10.1109/ichi48887.2020.9374399>

# Real life usage of TRN Model

The Temporal Recurrent Network (TRN) model has been proposed as a framework for online action detection and anticipation of the immediate future in videos (Xu et al., 2019). It aims to model greater temporal context by simultaneously performing these tasks. The TRN model allows for the learning and reasoning of temporal dependencies between video frames at multiple time scales (Zhou et al., 2018). This makes it an effective and interpretable network module for understanding and analyzing videos.

The TRN model has also been studied in the context of the thalamic reticular nucleus (TRN) in neuroscience research. The TRN is known to regulate thalamocortical interactions critical for sensory processing, attention, and cognition (Li et al., 2020). It has been found that the TRN plays a role in state-regulation in vivo and can affect behavior (Halassa et al., 2011). Dysfunction in the TRN has been implicated in neurodevelopmental disorders and has been associated with measurable deficits (Krol et al., 2018). Additionally, oxidative stress has been shown to lead to T-type calcium channel hypofunction in the TRN, which may be relevant to conditions like schizophrenia (Khoueiry et al., 2022).

The TRN model has also been applied in other fields. In computational biology, TRN models have been used to understand transcriptional regulatory networks and gene regulatory networks (He & Tan, 2016). These models have been constructed using genomic technologies and computational modeling techniques (He & Tan, 2016). In the field of mathematics education, the concept of "application and modeling" has been used to explain the relationship between real-life situations and mathematical concepts (Arseven, 2015). Machine learning models, including the TRN model, have been used for predicting energy consumption in buildings (Kim et al., 2016).

In summary, the TRN model has been used in various domains, including computer science, neuroscience, computational biology, mathematics education, and building energy prediction. It has been applied to tasks such as online action detection, understanding temporal dependencies in videos, studying the role of the TRN in behavior and neurodevelopmental disorders, and modeling transcriptional regulatory networks and gene regulatory networks. The TRN model offers a framework for analyzing and understanding complex temporal relationships and has the potential to advance research in these fields.

References:

Arseven, A. (2015). Mathematical modelling approach in mathematics education. Universal Journal of Educational Research, 3(12), 973-980. https://doi.org/10.13189/ujer.2015.031204

Halassa, M., Siegle, J., Ritt, J., Ting, J., Feng, G., & Moore, C. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. Nature Neuroscience, 14(9), 1118-1120. https://doi.org/10.1038/nn.2880

He, B. and Tan, K. (2016). Understanding transcriptional regulatory networks using computational models. Current Opinion in Genetics & Development, 37, 101-108. https://doi.org/10.1016/j.gde.2016.02.002

Khoueiry, C., Cabungcal, J., Rovó, Z., Fournier, M., Do, K., & Steullet, P. (2022). Developmental oxidative stress leads to t-type ca2+ channel hypofunction in thalamic reticular nucleus of mouse models pertinent to schizophrenia. Molecular Psychiatry, 27(4), 2042-2051. https://doi.org/10.1038/s41380-021-01425-2

Kim, Y., Ahn, K., & Park, C. (2016). Issues of application of machine learning models for virtual and real-life buildings. Sustainability, 8(6), 543. https://doi.org/10.3390/su8060543

Krol, A., Wimmer, R., Halassa, M., & Feng, G. (2018). Thalamic reticular dysfunction as a circuit endophenotype in neurodevelopmental disorders. Neuron, 98(2), 282-295. https://doi.org/10.1016/j.neuron.2018.03.021

Li, Y., López-Huerta, V., Adiconis, X., Levandowski, K., Choi, S., Simmons, S., … & Feng, G. (2020). Distinct subnetworks of the thalamic reticular nucleus. Nature, 583(7818), 819-824. https://doi.org/10.1038/s41586-020-2504-5

Xu, M., Gao, M., Chen, Y., Davis, L., & Crandall, D. (2019). Temporal recurrent networks for online action detection.. https://doi.org/10.1109/iccv.2019.00563

Zhou, B., Andonian, A., Oliva, A., & Torralba, A. (2018). Temporal relational reasoning in videos., 831-846. <https://doi.org/10.1007/978-3-030-01246-5_49>

# TRN model in neuroscience

Recent advances in genomic technologies and computational modeling have revolutionized our ability to construct models of transcriptional regulatory networks (TRNs) (He & Tan, 2016).

Transcranial random noise stimulation (tRNS) is a novel method of electrical stimulation that applies a random electrical oscillation spectrum over the motor cortex (Terney et al., 2008).

Based on the available references, we can discuss the usage of the TRN model in neuroscience. Recent advances in genomic technologies and computational modeling have greatly enhanced our understanding of transcriptional regulatory networks (He & Tan, 2016). These models have allowed us to study the complex interactions between genes and their regulatory elements, providing insights into the mechanisms underlying gene expression and cellular processes. Additionally, transcranial random noise stimulation (tRNS) has been used to increase human brain excitability, offering potential applications in neuroscience research (Terney et al., 2008).

References:

Arseven, A. (2015). Mathematical modelling approach in mathematics education. Universal Journal of Educational Research, 3(12), 973-980. https://doi.org/10.13189/ujer.2015.031204

Halassa, M., Siegle, J., Ritt, J., Ting, J., Feng, G., & Moore, C. (2011). Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. Nature Neuroscience, 14(9), 1118-1120. https://doi.org/10.1038/nn.2880

He, B. and Tan, K. (2016). Understanding transcriptional regulatory networks using computational models. Current Opinion in Genetics & Development, 37, 101-108. https://doi.org/10.1016/j.gde.2016.02.002

He, B. and Tan, K. (2016). Understanding transcriptional regulatory networks using computational models. Current Opinion in Genetics & Development, 37, 101-108. https://doi.org/10.1016/j.gde.2016.02.002

Khoueiry, C., Cabungcal, J., Rovó, Z., Fournier, M., Do, K., & Steullet, P. (2022). Developmental oxidative stress leads to t-type ca2+ channel hypofunction in thalamic reticular nucleus of mouse models pertinent to schizophrenia. Molecular Psychiatry, 27(4), 2042-2051. https://doi.org/10.1038/s41380-021-01425-2

Kim, Y., Ahn, K., & Park, C. (2016). Issues of application of machine learning models for virtual and real-life buildings. Sustainability, 8(6), 543. https://doi.org/10.3390/su8060543

Krol, A., Wimmer, R., Halassa, M., & Feng, G. (2018). Thalamic reticular dysfunction as a circuit endophenotype in neurodevelopmental disorders. Neuron, 98(2), 282-295. https://doi.org/10.1016/j.neuron.2018.03.021

Li, Y., López-Huerta, V., Adiconis, X., Levandowski, K., Choi, S., Simmons, S., … & Feng, G. (2020). Distinct subnetworks of the thalamic reticular nucleus. Nature, 583(7818), 819-824. https://doi.org/10.1038/s41586-020-2504-5

Terney, D., Chaieb, L., Moliadze, V., Antal, A., & Paulus, W. (2008). Increasing human brain excitability by transcranial high-frequency random noise stimulation. Journal of Neuroscience, 28(52), 14147-14155. https://doi.org/10.1523/jneurosci.4248-08.2008

Xu, M., Gao, M., Chen, Y., Davis, L., & Crandall, D. (2019). Temporal recurrent networks for online action detection.. https://doi.org/10.1109/iccv.2019.00563

Zhou, B., Andonian, A., Oliva, A., & Torralba, A. (2018). Temporal relational reasoning in videos., 831-846. <https://doi.org/10.1007/978-3-030-01246-5_49>

# Machine learning models, including the TRN model, have been used for predicting energy consumption in buildings

The TRN (Trees vs Neurons) model is a machine learning model that has been used for predicting energy consumption in buildings (Ahmad et al., 2017). This model utilizes both random forest and artificial neural network (ANN) algorithms to make high-resolution predictions of building energy consumption (Ahmad et al., 2017). The model takes into account various factors such as mean squared error, random forest, and categorical variables (Ahmad et al., 2017). By leveraging machine learning techniques, the TRN model can effectively analyze historical data and identify patterns and relationships that contribute to energy consumption in buildings.

The TRN model is part of a broader trend in the field of building energy prediction, where machine learning algorithms are increasingly being used to forecast energy consumption (Mounter et al., 2021). These models have shown promise in accurately predicting energy usage during the occupancy stage of buildings (Mounter et al., 2021). They have been compared to other machine learning techniques using the same datasets, demonstrating their ability to effectively predict building energy usage (Mounter et al., 2021).

In addition to the TRN model, other machine learning algorithms have also been applied to the prediction of energy consumption in buildings. For example, deep learning models have been used to predict building energy consumption, achieving accurate results (Wang, 2022). Extreme learning machine algorithms and support vector regression have also been employed to build energy consumption prediction models (Liu et al., 2023). These models leverage the power of machine learning to analyze data and make accurate predictions about energy usage in buildings.

Overall, machine learning models, including the TRN model, have proven to be effective tools for predicting energy consumption in buildings. By analyzing historical data and identifying patterns, these models can provide valuable insights into energy usage and help optimize energy efficiency in buildings.

References:

Ahmad, M., Mourshed, M., & Rezgui, Y. (2017). Trees vs neurons: comparison between random forest and ann for high-resolution prediction of building energy consumption. Energy and Buildings, 147, 77-89. https://doi.org/10.1016/j.enbuild.2017.04.038

Liu, R., Huang, L., & Yuan, C. (2023). Building energy consumption prediction based on vmd-cnn-bilstm-attention. Highlights in Science Engineering and Technology, 43, 185-197. https://doi.org/10.54097/hset.v43i.7419

Mounter, W., Ogwumike, C., Dawood, H., & Dawood, N. (2021). Machine learning and data segmentation for building energy use prediction—a comparative study. Energies, 14(18), 5947. https://doi.org/10.3390/en14185947

Wang, Y. (2022). Application of deep learning model in building energy consumption prediction. Computational Intelligence and Neuroscience, 2022, 1-9. <https://doi.org/10.1155/2022/4835259>

# video classification models

Video classification models play a crucial role in accurately categorizing and understanding the content of videos. Several studies have been conducted to develop effective models for video classification.

One approach is to design Convolutional Neural Network (CNN) architectures specifically for action recognition in videos. Wang et al. (2016) aim to discover the principles for designing effective ConvNet architectures for action recognition and learn these models with limited training samples. They propose Temporal Segment Networks, which achieve state-of-the-art performance on benchmark datasets.

Another popular model for image recognition, which can also be applied to video classification, is the Residual Network (ResNet) (He et al., 2015). (He et al., 2015) introduce deep residual learning, which allows for the training of very deep networks by using residual connections. This approach has been successful in various computer vision tasks, including image recognition.

Carreira & Zisserman (2017) address the challenge of limited videos in action classification datasets. They propose a new model and introduce the Kinetics dataset, which contains a large number of videos across multiple action categories. This dataset enables the identification of good video architectures and the evaluation of performance on a larger scale.

Deep learning models have also been applied to sports video classification. Guangyu (2022) proposes a sports video image classification method based on a deep learning coding model. This method extracts features from sports video images and utilizes a deep learning model for classification, achieving accurate sports video classification.

In addition to image-based approaches, text-based models have also been explored for video classification. Liu et al. (2021) propose an extended text combination classification model for short videos. This model compares favorably to existing methods and demonstrates accuracy and robustness in classifying short video titles.

Collaborative learning models have been developed to exploit the complementarity between static and motion information in videos. Peng et al. (2019) propose a two-stream collaborative learning model with spatial-temporal attention. This model achieves mutual guidance on static and motion information and adaptively learns the fusion weights of static and motion streams, leading to improved video classification performance.

Overall, video classification models have been developed using various deep learning architectures, including CNNs and ResNets. These models have been applied to different domains, such as action recognition, sports video classification, and short video classification. The performance of these models has been evaluated on benchmark datasets, demonstrating their effectiveness in accurately categorizing videos.

References:

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Guangyu, H. (2022). Analysis of sports video intelligent classification technology based on neural network algorithm and transfer learning. Computational Intelligence and Neuroscience, 2022, 1-10. https://doi.org/10.1155/2022/7474581

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.. https://doi.org/10.48550/arxiv.1512.03385

Liu, Y., Zhang, Y., Hu, H., Liu, X., Zhang, L., & Liu, R. (2021). An extended text combination classification model for short video based on albert. Journal of Sensors, 2021, 1-7. https://doi.org/10.1155/2021/8013337

Peng, Y., Zhao, Y., & Zhang, J. (2019). Two-stream collaborative learning with spatial-temporal attention for video classification. Ieee Transactions on Circuits and Systems for Video Technology, 29(3), 773-786. https://doi.org/10.1109/tcsvt.2018.2808685

Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., … & Gool, L. (2016). Temporal segment networks: towards good practices for deep action recognition., 20-36. <https://doi.org/10.1007/978-3-319-46484-8_2>

# what is video classification model?

Video classification models are algorithms or architectures designed to categorize and understand the content of videos. These models utilize various techniques, such as deep learning, to analyze and classify videos based on their visual features, temporal information, or textual content.

One widely used approach for video classification is the application of Convolutional Neural Networks (CNNs). Wang et al. (2016) propose Temporal Segment Networks, which combine sparse temporal sampling and video-level supervision to efficiently learn from action videos. This approach enables effective action recognition by utilizing the entire video.

Another important model for video classification is the Residual Network (ResNet) (He et al., 2015). introduce deep residual learning, which allows for the training of very deep networks by using residual connections. ResNet has been successfully applied to image recognition tasks and can also be adapted for video classification.

Transfer learning is another technique commonly employed in video classification models. Rafiq et al. (2020) focus on scene classification for sports video summarization using transfer learning. Scene classification plays a crucial role in video analysis and summarization, and the quality of scene classification greatly impacts the overall performance of these tasks.

In the context of sports video classification, Guangyu (2022) proposes a method based on deep learning coding models. This approach extracts features from sports video images and utilizes deep learning algorithms for accurate classification of sports videos.

Text-based models have also been explored for video classification. Liu et al. (2021) present an extended text combination classification model for short videos. This model leverages text information, such as video titles, to improve the accuracy and robustness of short video classification.

Additionally, spatial image polynomial decomposition has been utilized for video classification (Moubtahij et al., 2015). This approach involves modeling motion and texture features to describe actions in videos. The polynomial transform is applied to decompose the video and extract discriminative features for classification.

In summary, video classification models employ various techniques, including CNNs, ResNets, transfer learning, and text-based approaches, to accurately categorize and understand the content of videos. These models have been applied to different domains, such as action recognition, sports video classification, and short video classification, and have demonstrated their effectiveness in accurately classifying videos.

References:

Carreira, J. and Zisserman, A. (2017). Quo vadis, action recognition? a new model and the kinetics dataset.. https://doi.org/10.1109/cvpr.2017.502

Guangyu, H. (2022). Analysis of sports video intelligent classification technology based on neural network algorithm and transfer learning. Computational Intelligence and Neuroscience, 2022, 1-10. https://doi.org/10.1155/2022/7474581

Guangyu, H. (2022). Analysis of sports video intelligent classification technology based on neural network algorithm and transfer learning. Computational Intelligence and Neuroscience, 2022, 1-10. https://doi.org/10.1155/2022/7474581

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.. https://doi.org/10.48550/arxiv.1512.03385

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.. https://doi.org/10.48550/arxiv.1512.03385

Liu, Y., Zhang, Y., Hu, H., Liu, X., Zhang, L., & Liu, R. (2021). An extended text combination classification model for short video based on albert. Journal of Sensors, 2021, 1-7. https://doi.org/10.1155/2021/8013337

Liu, Y., Zhang, Y., Hu, H., Liu, X., Zhang, L., & Liu, R. (2021). An extended text combination classification model for short video based on albert. Journal of Sensors, 2021, 1-7. https://doi.org/10.1155/2021/8013337

Moubtahij, R., Augereau, B., Tairi, H., & Fernandez-Maloigne, C. (2015). Spatial image polynomial decomposition with application to video classification. Journal of Electronic Imaging, 24(6), 061114. https://doi.org/10.1117/1.jei.24.6.061114

Peng, Y., Zhao, Y., & Zhang, J. (2019). Two-stream collaborative learning with spatial-temporal attention for video classification. Ieee Transactions on Circuits and Systems for Video Technology, 29(3), 773-786. https://doi.org/10.1109/tcsvt.2018.2808685

Rafiq, M., Rafiq, G., Agyeman, R., Choi, G., & Jin, S. (2020). Scene classification for sports video summarization using transfer learning. Sensors, 20(6), 1702. https://doi.org/10.3390/s20061702

Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., … & Gool, L. (2016). Temporal segment networks: towards good practices for deep action recognition., 20-36. https://doi.org/10.1007/978-3-319-46484-8\_2

Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., … & Gool, L. (2016). Temporal segment networks: towards good practices for deep action recognition., 20-36. https://doi.org/10.1007/978-3-319-46484-8\_2

# what fields does video classification models use besides the deep learning?

Video classification models utilize various fields in addition to deep learning. Some of these fields include computer science, artificial intelligence, machine learning, convolutional neural networks (CNNs), support vector machines (SVMs), and data mining techniques.

Computer science plays a crucial role in video classification, as it provides the foundation for developing and implementing algorithms and models. Deep learning, a subfield of machine learning, is a popular approach used in video classification models (Rehman & Belhaouari, 2021). Deep learning techniques, such as CNNs, have been extensively studied and applied in video classification tasks (He et al., 2015; Funke et al., 2019; Minhas et al., 2019). CNNs automatically learn discriminant features for video classification and leverage the abundance of large amounts of online videos (Chatbri et al., 2017).

Artificial intelligence is another field that contributes to video classification models. It involves the development of intelligent algorithms and systems that can mimic human cognitive abilities, such as perception and decision-making. Deep learning-based approaches, which fall under the umbrella of artificial intelligence, have shown promising results in video classification (Pandey et al., 2021).

Machine learning techniques, including SVMs, are also utilized in video classification models (Mohanan, 2017; Gokila et al., 2019). SVMs are supervised learning models that can classify data into different categories based on training examples. They have been applied to classify different events in field sports videos (Minhas et al., 2019).

Data mining techniques are employed to extract useful information and patterns from large datasets. These techniques can be used to analyze video content and extract relevant features for classification (Shambharkar & Doja, 2015). They have been used in the automatic classification of movie trailers (Shambharkar & Doja, 2015).

In summary, video classification models draw upon various fields such as computer science, artificial intelligence, machine learning, convolutional neural networks, support vector machines, and data mining techniques. These fields provide the necessary tools and techniques to develop effective and accurate video classification models.

References:

Chatbri, H., Oliveira, M., McGuinness, K., Little, S., Kameyama, K., Kwan, P., … & O'Connor, N. (2017). Educational video classification by using a transcript to image transform and supervised learning.. https://doi.org/10.1109/ipta.2017.8853988

Funke, I., Mees, S., Weitz, J., & Speidel, S. (2019). Video-based surgical skill assessment using 3d convolutional neural networks. International Journal of Computer Assisted Radiology and Surgery, 14(7), 1217-1225. https://doi.org/10.1007/s11548-019-01995-1

Gokila, K., Jananie, K., C, M., & M, G. (2019). Content based video classification using svm. International Journal of Recent Trends in Engineering and Research, 5(Special Issue 1), 468-477. https://doi.org/10.23883/ijrter.conf.20190322.060.iwtaz

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition.. https://doi.org/10.48550/arxiv.1512.03385

Minhas, R., Javed, A., Irtaza, A., Mahmood, M., & Joo, Y. (2019). Shot classification of field sports videos using alexnet convolutional neural network. Applied Sciences, 9(3), 483. https://doi.org/10.3390/app9030483

Mohanan, S. (2017). Sports video categorization by multiclass svm using higher order spectra features. International Journal of Advances in Signal and Image Sciences, 3(2), 27. https://doi.org/10.29284/ijasis.3.2.2017.27-33

Pandey, R., Chikhale, Y., Verma, R., & Patil, D. (2021). Deep learning based human action recognition. Itm Web of Conferences, 40, 03014. https://doi.org/10.1051/itmconf/20214003014

Rehman, A. and Belhaouari, S. (2021). Deep learning for video classification: a review.. https://doi.org/10.36227/techrxiv.15172920.v1

Shambharkar, P. and Doja, M. (2015). Automatic classification of movie trailers using data mining techniques: a review.. <https://doi.org/10.1109/ccaa.2015.7148349>

# Deep learning in the context of video classification models

Deep learning has emerged as a powerful technique for video classification models. It involves the use of deep neural networks, specifically convolutional neural networks (CNNs), to extract high-level features from video frames and make predictions based on those features. Several studies have explored different aspects of deep learning in the context of video classification.

One important aspect is the design of effective ConvNet architectures for action recognition in videos. Wang et al. (2016) aim to discover the principles for designing such architectures and learn these models given limited training samples. They emphasize the importance of deep visual representations in computer vision and provide two best-performing ConvNet models for further research.

Another aspect is the use of transfer learning in video classification. Rafiq et al. (2020) propose a deep-learning-based shot classification model for sports video summarization using transfer learning. They achieve good accuracy over different shot classes by leveraging pre-trained models and transfer learning techniques.

The temporal aspect of videos is also crucial for accurate classification. Kiziltepe et al. (2021) employ recurrent neural networks (RNNs) to gather temporal information and improve video classification performance. They highlight the power of RNNs in sequence learning and their ability to capture temporal dependencies in videos.

In addition to action recognition, deep learning has been applied to other types of video classification tasks. For example, Tuan et al. (2020) present a deep learning-based multimodal system for classifying daily life videos. They demonstrate the effectiveness of deep learning techniques in classifying videos based on their content.

Furthermore, deep learning has been used in combination with other techniques for video classification. Wu et al. (2015) propose a hybrid deep learning framework that models static spatial information, short-term motion, and long-term temporal clues in videos. This framework combines the strengths of different deep learning architectures to improve video classification performance.

Overall, deep learning has shown great potential in the context of video classification models. It enables the extraction of high-level features from video frames, captures temporal dependencies, and achieves high accuracy in various video classification tasks. By leveraging transfer learning and combining different deep learning architectures, researchers have made significant advancements in this field.

References:

Kiziltepe, R., Gan, J., & Escobar, J. (2021). A novel keyframe extraction method for video classification using deep neural networks. Neural Computing and Applications. https://doi.org/10.1007/s00521-021-06322-x

Rafiq, M., Rafiq, G., Agyeman, R., Choi, G., & Jin, S. (2020). Scene classification for sports video summarization using transfer learning. Sensors, 20(6), 1702. https://doi.org/10.3390/s20061702

Tuan, P., Nguyen, A., Ngo, T., Le, D., Le, K., Nguyen, T., … & Le, H. (2020). Proposed smart university model as a sustainable living lab for university digital transformation.. https://doi.org/10.1109/gtsd50082.2020.9303086

Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., … & Gool, L. (2016). Temporal segment networks: towards good practices for deep action recognition., 20-36. https://doi.org/10.1007/978-3-319-46484-8\_2

Wu, Z., Wang, X., Jiang, Y., Hao, Y., & Xue, X. (2015). Modeling spatial-temporal clues in a hybrid deep learning framework for video classification.. <https://doi.org/10.1145/2733373.2806222>

# What are real time usages of deep learning in video classification models?

Deep learning has found various real-time applications in video classification models. One such application is in the field of video surveillance, where deep learning technologies have transformed traditional systems into smart video surveillance systems that automate human behavior classification (AlQaralleh et al., 2022). Another application is in vehicle detection and classification, where a real-time system utilizes a deep learning tool called You Only Look Once (YOLO) for object detection and classification (Kwan et al., 2020). Deep learning has also been applied in sports analysis, such as basketball, where it is used for real-time analysis of basketball sports data (Peng, 2021). In the field of computer vision, deep learning has been used for real-time object detection and recognition, such as car detection using unmanned aerial vehicles (Benjdira et al., 2019). Additionally, deep learning has been utilized in real-time video processing, such as real-time yoga pose recognition and video processing for yoga exercises (Sharma et al., 2021; Yadav et al., 2019).

Furthermore, deep learning has been applied in the domain of video forgery detection, where it has shown success in detecting duplicate frames in videos (Munawar & Noreen, 2021). It has also been used for video aesthetic quality assessment, where deep multimodality learning has been employed to assess the aesthetic quality of UAV videos (Kuang et al., 2020). In the medical field, deep learning has been used for real-time segmentation of kidney stones in endoscopic video feeds (Stoebner, 2022). Moreover, deep learning has been applied in the context of underwater videos, where it is used for fish segmentation and classification (Jahanbakht et al., 2022).

Overall, deep learning has proven to be a powerful tool in real-time video classification models, enabling applications in various domains such as surveillance, sports analysis, computer vision, forgery detection, aesthetic assessment, medical imaging, and underwater video processing.

References:

AlQaralleh, E., Aldhaban, F., Nasseif, H., Alksasbeh, M., & Alqaralleh, B. (2022). Smart deep learning based human behaviour classification for video surveillance. Computers Materials & Continua, 72(3), 5593-5605. https://doi.org/10.32604/cmc.2022.026666

Benjdira, B., Khursheed, T., Koubâa, A., Ammar, A., & Ouni, K. (2019). Car detection using unmanned aerial vehicles: comparison between faster r-cnn and yolov3.. https://doi.org/10.1109/uvs.2019.8658300

Jahanbakht, M., Wang, X., Waltham, N., & Azghadi, M. (2022). Distributed deep learning and energy-efficient real-time image processing at the edge for fish segmentation in underwater videos. Ieee Access, 10, 117796-117807. https://doi.org/10.1109/access.2022.3202975

Kuang, Q., Jin, X., Zhao, Q., & Zhou, B. (2020). Deep multimodality learning for uav video aesthetic quality assessment. Ieee Transactions on Multimedia, 22(10), 2623-2634. https://doi.org/10.1109/tmm.2019.2960656

Kwan, C., Gribben, D., Chou, B., Budavari, B., Larkin, J., Rangamani, A., … & Etienne-Cummings, R. (2020). Real-time and deep learning based vehicle detection and classification using pixel-wise code exposure measurements. Electronics, 9(6), 1014. https://doi.org/10.3390/electronics9061014

Munawar, M. and Noreen, I. (2021). Duplicate frame video forgery detection using siamese-based rnn. Intelligent Automation & Soft Computing, 29(3), 927-937. https://doi.org/10.32604/iasc.2021.018854

Peng, Y. (2021). Real-time analysis of basketball sports data based on deep learning. Complexity, 2021, 1-11. https://doi.org/10.1155/2021/9142697

Sharma, V., Gupta, M., Kumar, A., & Mishra, D. (2021). Video processing using deep learning techniques: a systematic literature review. Ieee Access, 9, 139489-139507. https://doi.org/10.1109/access.2021.3118541

Stoebner, Z. (2022). Segmentation of kidney stones in endoscopic video feeds.. https://doi.org/10.48550/arxiv.2204.14175

Yadav, S., Singh, A., Gupta, A., & Raheja, J. (2019). Real-time yoga recognition using deep learning. Neural Computing and Applications, 31(12), 9349-9361. <https://doi.org/10.1007/s00521-019-04232-7>

# What are the usages of SVM in video classification models?

SVM (Support Vector Machine) is a supervised learning algorithm that is commonly used for classification and regression tasks (Alaoui et al., 2019). In the context of video classification models, SVM has several usages and applications.

One of the main usages of SVM in video classification is for sports video categorization. SVM has been successfully applied to classify sports videos based on their content and genre (Chattopadhyay & Maurya, 2013; Ansari & Vasishtha, 2016; Li et al., 2009). For example, a study used SVM to classify sports teaching videos based on their content using video and audio dual-mode characteristics (Zhang & Li, 2021). Another study utilized SVM with Higher Order Spectra Features (HOSF) for feature extraction and multiclass SVM classifier for the classification of sports videos (Mohanan, 2017). SVM has also been used for the classification of specific sports genres, such as Wushu videos (Hu & Zhai, 2009).

SVM is also used for action-based video classification. It has been employed to classify action recognition (AR) in videos (kaur & Doegar, 2019). SVM classifier is used to classify different actions based on heterogeneous feature extraction (kaur & Doegar, 2019).

Furthermore, SVM has been used in the classification of other types of videos as well. For instance, SVM has been used for fall detection in videos of elderly people by analyzing the variation of key points of the human skeleton (Alaoui et al., 2019). SVM has also been used for the classification of colonoscopy images in video-specific SVMs (Manivannan et al., 2014).

In addition to video classification, SVM has been used in various other domains such as audio signal classification (Dhanalakshmi et al., 2009), advertisement link classification (Xu et al., 2017), and moving target classification in video sequences (Kong & Wang, 2010). SVM has also been combined with other techniques like convolutional neural networks (CNN) for lung sound classification Aykanat et al. (2017) and memetic algorithm for imbalanced classification (Wu et al., 2011).

Overall, SVM is a versatile and widely used algorithm in video classification models. It has been applied to various domains and has shown good performance in classifying different types of videos based on their content, genre, and actions. SVM's ability to handle multi-class classification problems and its effectiveness in feature extraction make it a popular choice for video classification tasks.

References:

Alaoui, A., Fkihi, S., & Thami, R. (2019). Fall detection for elderly people using the variation of key points of human skeleton. Ieee Access, 7, 154786-154795. https://doi.org/10.1109/access.2019.2946522

Ansari, M. and Vasishtha, H. (2016). Content based video retrieval systems performance based on multiple features and multiple frames using svm. International Journal of Advanced Computer Science and Applications, 7(8). https://doi.org/10.14569/ijacsa.2016.070815

Aykanat, M., Kılıç, Ö., Kurt, B., & Saryal, S. (2017). Classification of lung sounds using convolutional neural networks. Eurasip Journal on Image and Video Processing, 2017(1). https://doi.org/10.1186/s13640-017-0213-2

Chattopadhyay, C. and Maurya, A. (2013). Genre-specific modeling of visual features for efficient content based video shot classification and retrieval. International Journal of Multimedia Information Retrieval, 2(4), 289-297. https://doi.org/10.1007/s13735-013-0034-8

Dhanalakshmi, P., Palanivel, S., & Ramalingam, V. (2009). Classification of audio signals using svm and rbfnn. Expert Systems With Applications, 36(3), 6069-6075. https://doi.org/10.1016/j.eswa.2008.06.126

Hu, Y. and Zhai, G. (2009). Automatic classification system of wushu video based on svm.. https://doi.org/10.1109/fskd.2009.438

Kong, Y. and Wang, L. (2010). Moving target classification in video sequences based on features combination and svm.. https://doi.org/10.1109/cise.2010.5676969

Li, L., Zhang, N., Duan, L., Huang, Q., Du, J., & Guan, L. (2009). Automatic sports genre categorization and view-type classification over large-scale dataset.. https://doi.org/10.1145/1631272.1631380

Manivannan, S., Wang, R., Trujillo, M., Hoyos, J., & Trucco, E. (2014). Video-specific svms for colonoscopy image classification., 11-21. https://doi.org/10.1007/978-3-319-13410-9\_2

Mohanan, S. (2017). Sports video categorization by multiclass svm using higher order spectra features. International Journal of Advances in Signal and Image Sciences, 3(2), 27. https://doi.org/10.29284/ijasis.3.2.2017.27-33

Wu, M., Watada, J., Ibrahim, Z., & Khalid, M. (2011). Building a memetic algorithm based support vector machine for imbalaced classification.. https://doi.org/10.1109/icgec.2011.95

Xu, J., Chang, P., & Zhang, Y. (2017). Discovering advertisement links by using url text.. https://doi.org/10.23977/iccsc.2017.1020

Zhang, M. and Li, W. (2021). An automatic classification method of sports teaching video using support vector machine. Scientific Programming, 2021, 1-8. https://doi.org/10.1155/2021/4728584

kaur, C. and Doegar, A. (2019). Classification of action based video using heterogeneous feature extraction and svm. International Journal of Innovative Technology and Exploring Engineering, 8(11), 1887-1892. <https://doi.org/10.35940/ijitee.k2089.0981119>

# What is SVM?

SVM (Support Vector Machine) is a machine learning algorithm commonly used for classification and regression tasks (Cortes & Vapnik, 1995). It is a versatile algorithm that has been applied in various fields, including computer science, biology, and chemistry.

In the context of machine learning, SVM is known for its ability to handle both linear and non-linear classification problems (Cortes & Vapnik, 1995). It works by finding an optimal hyperplane that separates the data points of different classes with the largest margin. This hyperplane is determined by support vectors, which are the data points closest to the decision boundary (Cortes & Vapnik, 1995). SVM aims to maximize the margin between the support vectors, which leads to better generalization and robustness of the model.

SVM has been extensively used in image classification tasks. It is known for its ability to generalize well even with limited training samples (Rapiee et al., 2022). SVM classifiers have been applied in remote sensing technology for mapping and classifying different land cover types (Rapiee et al., 2022). The ability of SVM to handle high-dimensional data and its effectiveness in generalization make it a popular choice for image classification tasks.

In addition to image classification, SVM has also been used in other domains. In the field of chemistry, SVM has been employed for compound classification and property prediction (Siemers & Bajorath, 2023). It has been compared with other machine learning algorithms, such as random forest, to analyze the differences in learning characteristics (Siemers & Bajorath, 2023). SVM has also been used for activity prediction in chemical compounds, and visualization techniques have been developed to interpret the SVM predictions (Balfer & Bajorath, 2015).

Overall, SVM is a versatile machine learning algorithm that has found applications in various fields. Its ability to handle both linear and non-linear classification problems, along with its generalization capabilities, make it a popular choice for classification and regression tasks. SVM has been successfully applied in image classification, compound classification, and other domains, showcasing its effectiveness and versatility in different contexts.

References:

Alaoui, A., Fkihi, S., & Thami, R. (2019). Fall detection for elderly people using the variation of key points of human skeleton. Ieee Access, 7, 154786-154795. https://doi.org/10.1109/access.2019.2946522

Ansari, M. and Vasishtha, H. (2016). Content based video retrieval systems performance based on multiple features and multiple frames using svm. International Journal of Advanced Computer Science and Applications, 7(8). https://doi.org/10.14569/ijacsa.2016.070815

Aykanat, M., Kılıç, Ö., Kurt, B., & Saryal, S. (2017). Classification of lung sounds using convolutional neural networks. Eurasip Journal on Image and Video Processing, 2017(1). https://doi.org/10.1186/s13640-017-0213-2

Balfer, J. and Bajorath, J. (2015). Visualization and interpretation of support vector machine activity predictions. Journal of Chemical Information and Modeling, 55(6), 1136-1147. https://doi.org/10.1021/acs.jcim.5b00175

Chattopadhyay, C. and Maurya, A. (2013). Genre-specific modeling of visual features for efficient content based video shot classification and retrieval. International Journal of Multimedia Information Retrieval, 2(4), 289-297. https://doi.org/10.1007/s13735-013-0034-8

Cortes, C. and Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. https://doi.org/10.1007/bf00994018

Dhanalakshmi, P., Palanivel, S., & Ramalingam, V. (2009). Classification of audio signals using svm and rbfnn. Expert Systems With Applications, 36(3), 6069-6075. https://doi.org/10.1016/j.eswa.2008.06.126

Hu, Y. and Zhai, G. (2009). Automatic classification system of wushu video based on svm.. https://doi.org/10.1109/fskd.2009.438

Kong, Y. and Wang, L. (2010). Moving target classification in video sequences based on features combination and svm.. https://doi.org/10.1109/cise.2010.5676969

Li, L., Zhang, N., Duan, L., Huang, Q., Du, J., & Guan, L. (2009). Automatic sports genre categorization and view-type classification over large-scale dataset.. https://doi.org/10.1145/1631272.1631380

Manivannan, S., Wang, R., Trujillo, M., Hoyos, J., & Trucco, E. (2014). Video-specific svms for colonoscopy image classification., 11-21. https://doi.org/10.1007/978-3-319-13410-9\_2

Mohanan, S. (2017). Sports video categorization by multiclass svm using higher order spectra features. International Journal of Advances in Signal and Image Sciences, 3(2), 27. https://doi.org/10.29284/ijasis.3.2.2017.27-33

Rapiee, D., Kalu, M., & Suhaili, A. (2022). Mapping of shorea kunstleri king in mixed dipterocarp forest using remote sensing technology. Iop Conference Series Earth and Environmental Science, 1102(1), 012009. https://doi.org/10.1088/1755-1315/1102/1/012009

Siemers, F. and Bajorath, J. (2023). Differences in learning characteristics between support vector machine and random forest models for compound classification revealed by shapley value analysis. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-33215-x

Wu, M., Watada, J., Ibrahim, Z., & Khalid, M. (2011). Building a memetic algorithm based support vector machine for imbalaced classification.. https://doi.org/10.1109/icgec.2011.95

Xu, J., Chang, P., & Zhang, Y. (2017). Discovering advertisement links by using url text.. https://doi.org/10.23977/iccsc.2017.1020

Zhang, M. and Li, W. (2021). An automatic classification method of sports teaching video using support vector machine. Scientific Programming, 2021, 1-8. https://doi.org/10.1155/2021/4728584

kaur, C. and Doegar, A. (2019). Classification of action based video using heterogeneous feature extraction and svm. International Journal of Innovative Technology and Exploring Engineering, 8(11), 1887-1892. <https://doi.org/10.35940/ijitee.k2089.0981119>

# data mining with video classification models

Video classification is the task of categorizing videos into different classes or labels based on their content (Simonyan, 2014). It is an important area of research in computer vision and data mining (Simonyan, 2014). Several studies have been conducted to develop models and techniques for video classification.

One approach to video classification is using deep convolutional neural networks (CNNs). Simonyan (2014) proposed very deep CNNs for large-scale image recognition, which formed the basis for their team's submission to the ImageNet Challenge 2014 (Simonyan, 2014; . This approach achieved high accuracy in both localization and classification tasks (Simonyan, 2014). Similarly, Szegedy et al. (2016) rethought the inception architecture for computer vision and achieved substantial gains in classification accuracy (Szegedy et al., 2016). These studies demonstrate the effectiveness of deep CNNs for video classification.

Another approach to video classification is based on data mining techniques. Nguyen et al. (2022) proposed a novel representation of dynamic texture data for video classification using data mining (Nguyen et al., 2022). They used probabilistic motifs and data mining algorithms to classify videos based on their content (Nguyen et al., 2022). This approach shows the potential of data mining in video classification.

Text-based video content classification is another area of research in video classification. Huang et al. (2010) used feature-based classification techniques, such as C4.5, Naïve Bayes, and Support Vector Machine, to classify videos based on their text content (Huang et al., 2010). Liu et al. (2021) proposed an extended text combination classification model for short videos based on the ALBERT model (Liu et al., 2021). These studies highlight the importance of text analysis in video classification.

Educational data mining is also relevant to video classification. Tegegne & Alemu (2018) discussed the application of data mining techniques for analyzing students' academic performance in Ethiopian universities (Tegegne & Alemu, 2018). Classification is one of the commonly applied data mining techniques in educational data mining (Tegegne & Alemu, 2018). This demonstrates the potential of data mining in video classification for educational purposes (Tegegne & Alemu, 2018).

In summary, video classification can be approached using deep CNNs, data mining techniques, text analysis, and educational data mining (Simonyan, 2014; Szegedy et al., 2016; Nguyen et al., 2022; Huang et al., 2010; Liu et al., 2021; Tegegne & Alemu, 2018). These approaches have been shown to be effective in classifying videos based on their content (Simonyan, 2014; Szegedy et al., 2016; Nguyen et al., 2022; Huang et al., 2010; Liu et al., 2021; Tegegne & Alemu, 2018). Further research can explore the combination of these approaches to improve the accuracy and efficiency of video classification models.

References:

Huang, C., Fu, T., & Chen, H. (2010). Text‐based video content classification for online video‐sharing sites. Journal of the American Society for Information Science and Technology, 61(5), 891-906. https://doi.org/10.1002/asi.21291

Liu, Y., Zhang, Y., Hu, H., Liu, X., Zhang, L., & Liu, R. (2021). An extended text combination classification model for short video based on albert. Journal of Sensors, 2021, 1-7. https://doi.org/10.1155/2021/8013337

Nguyen, L., Mille, J., Li, D., Conte, D., & Ragot, N. (2022). Efficient dynamic texture classification with probabilistic motifs.. https://doi.org/10.1109/icpr56361.2022.9956557

Simonyan, K. (2014). Very deep convolutional networks for large-scale image recognition.. https://doi.org/10.48550/arxiv.1409.1556

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision.. https://doi.org/10.1109/cvpr.2016.308

Tegegne, A. and Alemu, T. (2018). Educational data mining for students’ academic performance analysis in selected ethiopian universities. Information Impact Journal of Information and Knowledge Management, 9(2), 1-15. https://doi.org/10.4314/iijikm.v9i2.1