VISUALIZING COALITION FORMATION: FROM HEDONIC GAMES TO IMAGE SEGMENTATION

Pedro H. de Paula França
UFRJ
pedro.franca@ufrj.br

Lucas Lopes Felipe
UFRJ
lucaslopesf2@gmail.com

Daniel Sadoc Menasché
UFRJ
sadoc@ic.ufrj.br

ABSTRACT

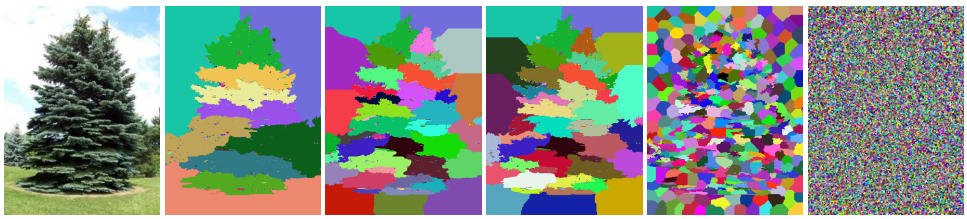
We propose image segmentation as a visual diagnostic testbed for coalition formation in hedonic games. Modeling pixels as agents on a graph, we study how a granularization parameter shapes equilibrium fragmentation and boundary structure. On the Weizmann single-object benchmark, we relate multi-coalition equilibria to binary protocols by measuring whether the converged coalitions overlap with a foreground ground-truth. We observe transitions from cohesive to fragmented yet recoverable equilibria, and finally to intrinsic failure under excessive fragmentation. Our core contribution links multi-agent systems with image segmentation by quantifying the impact of mechanism design parameters on equilibrium structures.

1 INTRODUCTION

Coalition formation is a common emergent behavior in multi-agent systems and mechanism design: agents optimize individual utilities under interaction constraints, and equilibria appear as partitions of the population (Bogomolnaia and Jackson, 2002). We leverage intuitive image segmentation to interpret a multi-agent coalition mechanism, providing a quantitative measure of a resolution parameter that controls coalition granularity. Image segmentation is an essential task across a range of scientific and technological domains (Mittal et al., 2022).

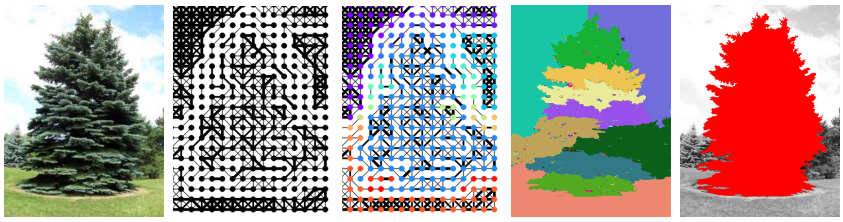
The *resolution* parameter γ modulates coalition granularity (Traag et al., 2011), fundamentally reshaping the equilibrium structure of the system (Felipe et al., 2025). By adjusting γ, the system can span the full spectrum of partitions, ranging from a single grand coalition as $\gamma \to 0$ to a set of individual singletons as $\gamma \to 1$. Figure 1 illustrates how image segmentation visualizes this transition from smaller to larger values of γ (Fig. 1(b) to Fig. 1(f)), mapping different partition equilibria to varying levels of spatial granularity. Indeed, a central challenge for mechanism designers is identifying the specific value of γ that yields meaningful equilibria for a given real-world application.

We bridge image segmentation and the proposed coalition mechanism by representing the image as a graph where pixels act as nodes. In this framework, the primary distinction between methods lies in the edge construction, which encodes pairwise similarities according to specific schemes (Shi and Malik, 1997; Galun et al., 2003; Alpert et al., 2007). This formulation effectively casts the segmentation task as a network clustering or community detection problem (Avrachenkov et al., 2018).



(a) (b) (c) (d) (e) (f)

Figure 1: Segmentation of the same image for increasing values of γ. (a) Original image. (b) $\gamma = 10^{-6}$. (c) $\gamma = 8 \times 10^{-6}$. (d) $\gamma = 10^{-5}$. (e) $\gamma = 5 \times 10^{-4}$. (f) $\gamma = 2.5 \times 10^{-1}$.



(a) (b) (c) (d) (e)

Figure 2: Pipeline: (a) image → (b) graph construction (downsampled,[1] see Section 4) → (c) equilibrium partition (Section 2) → (d) segmented image → (e) ground truth evaluation (Section 3).

Our key contribution lies in using the pipeline in Figure 2 to bridge the gap between multi-agent systems and intuitive image segmentation tasks, so as to quantify the impact of the resolution parameter on equilibrium structures. At the core of the pipeline lies the coalition mechanism (Fig. 2(c)), which determines the specific cluster assignment for each pixel; these assignments then serve as labels to reconstruct the segmented image (Fig. 2(d)). In Section 2, we detail the mechanism based on hedonic games (Felipe et al., 2025). Section 3 then evaluates the overlap between the equilibrium partitions produced by the mechanism and the ground-truth foreground coalition (Fig. 2(e)). Section 5 concludes with a summary of our contributions and future work.[1]

2 Hedonic Mechanism for Coalition Formation

A hedonic game is a naturally decentralized coalition formation game in which an agent's preference depends strictly on the composition of its own coalition (Aziz and Savani, 2016). A potential game allows unilateral improvements by individual agents to be aligned with the maximization of a global potential function, where local dynamics can be interpreted as hill-climbing a single objective (Monderer and Shapley, 1996).

Felipe et al. (2025) modeled the Constant Potts Model (CPM) (Traag et al., 2011), a quality function known for being a *resolution-limit-free* method (Fortunato and Barthelemy, 2007), as an *additively separable potential hedonic game*. The CPM is decomposed into individual agent utilities (hedonic potentials). This frames the mechanism as a non-cooperative game while remaining aligned with a global quality function, ensuring that selfish agents' decisions yield stable, cooperative coalitions.

Potential. Each player seeks to belong to the community that maximizes its utility, in which better balances attraction to well-connected neighbors against a penalty for joining large, weakly connected coalitions. For a node v and community C, we define

$$\text{Potential}_v^\gamma(C) = w(v, C) - \gamma(|C| - 1) \tag{1}$$

where $w(v, C) = \sum_{u \in C \setminus \{v\}} w_{vu}$ is the total edge weight from v to the vertices of C, and $|C|$ denotes the size of the candidate community C *after* assigning v to it (i.e., the community size used to evaluate the hypothetical move). The parameter $\gamma \in [0, 1]$ controls the trade-off between cohesion and coalition size: small γ favors larger cohesive regions, while larger γ penalizes large communities and promotes fragmentation into smaller coalitions (Figure 1).

Equilibrium. An *equilibrium* is a stable partition $\pi = \{C_1, \ldots, C_K\}$ where no agent has an incentive to change their community. This state is reached when every node is *locally stable*, meaning its current community C_{σ_v} provides a potential greater than or equal to any other community C_k:

$$\text{Potential}_v^\gamma(C_{\sigma_v}) \geq \text{Potential}_v^\gamma(C_k) \quad \forall k \neq \sigma_v. \tag{2}$$

A partition is considered a solution to the community detection problem if it satisfies two conditions:

- **Internal Stability:** No node wants to leave its community, e.g., to create a new community.
- **External Stability:** No node wants to join a different existing community.

[1] We detail our graph construction in Section 4. While it is impactful for performance, our mechanism is agnostic to the specific construction method as long as the input is a weighted graph; thus, evaluating alternatives remains future work. Figs. 2(b) and 2(c) are downsampled versions of Fig. 2(a) for visual clarity.

Algorithm 1: Hedonic optimization: The algorithm terminates at a locally stable equilibrium.

Input : Weighted Graph $G = (V, E, w)$, resolution γ, partition $\pi^{(0)}$
Output: Equilibrium partition $\pi = \{C_1, \ldots, C_K\}$

1 **if** $\pi^{(0)}$ *not provided*
2 | Assign each node to its own community ▷ `Initialize a singleton partition`
3 **change** $\leftarrow$ `true` ▷ `Activate the main improvement loop`
4 **while** ***change***
5 | **change** $\leftarrow$ `false` ▷ `Clear the change flag`
6 | **foreach** *node* $v \in V$
7 | | $\sigma_v \leftarrow$ current community of pixel v ▷ `Retrieve current community` σ_v
8 | | $\sigma^\star \leftarrow \arg\max_C \text{Potential}_v^\gamma(C)$ ▷ `Get community maximizing potential`
9 | | **if** $\text{Potential}_v^\gamma(\sigma^\star) > \text{Potential}_v^\gamma(\sigma_v)$
10 | | | Move v to $\sigma^\star$ ▷ `Reassign the node's community`
11 | | | **change** $\leftarrow$ `true` ▷ `Mark change occurred to force another pass`

In equilibrium, each node v is assigned to a community σ_v that maximizes its individual utility $\sigma_v \in \arg\max_k \text{Potential}_v^\gamma(C_k)$. Movement only occurs if a change offers a **strictly higher** utility. If multiple communities provide the same maximum utility, the node remains in its current community. Algorithm 1 outlines the hedonic game mechanism,[2] which relies on a potential function bounded by $2|V|^2$. Given a rational resolution parameter $\gamma = b/\kappa$, every improving move increases the potential by at least $1/\kappa$, guaranteeing convergence in $O(\kappa \cdot |V|^2)$ steps (Felipe et al., 2025).

3 RESULTS

To quantify how γ shapes equilibrium partitions and yields meaningful segmentations, we evaluate the converged partitions using two post-hoc accuracy metrics inspired by Alpert et al. (2007). For each image, we select one binary ground-truth (GT), denoted by Y, where $Y \in \{0, 1\}^{H \times W}$:

- $\text{F}_1^{\textbf{single}}$ **(Dominant-Coalition Accuracy):** Evaluates the F_1 score of the *single* community $C_k \in \pi$ that maximizes the F_1 score relative to Y;
- $\text{F}_1^{\textbf{union}}$ **(Recoverable-Union Accuracy):** Evaluates the F_1 score of the *subset* of communities $S \subseteq \{C_1, \ldots, C_K\}$ whose union maximizes the F_1 score relative to Y.

F_1 is the harmonic mean of precision and recall (Christen et al., 2023). Specifically, $\text{F}_1^{\text{single}}$ measures the emergence of a cohesive foreground, while $\text{F}_1^{\text{union}}$ assesses whether the foreground object, even if fragmented, remains "recoverable" from the equilibrium structure. In practice, throughout the remainder of the paper, when we refer to F_1^{union} we mean a proxy for F_1^{union} computed using the heuristic procedure described in the full version of this work (França et al., 2026).[3]

We use the *Weizmann Segmentation Evaluation Database* (Alpert et al., 2007) to obtain images with known ground truth, which is used exclusively at evaluation time. Note that we restrict our analysis to the single-object subset (100 natural images, 3 human ground-truth masks per image) and leave the two-object subset for future work. Experiments can be reproduced at: `https://github.com/henriquepedro1991/hedonic-games-to-image-segmentation`.

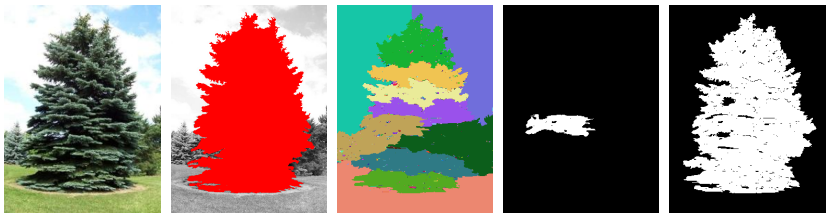
Resolution Design. We extend the approach of Avrachenkov et al. (2018)—setting the resolution as the graph's edge density—by proposing a scaling approach to evaluate varying density ratios:

$$\gamma = \frac{\text{density}(G)}{c}, \qquad \text{density}(G) = \frac{2|E|}{|V|(|V|-1)},$$

where c is a fixed constant. This normalization allows γ to act as a consistent local decision threshold across graphs of varying sparsity.

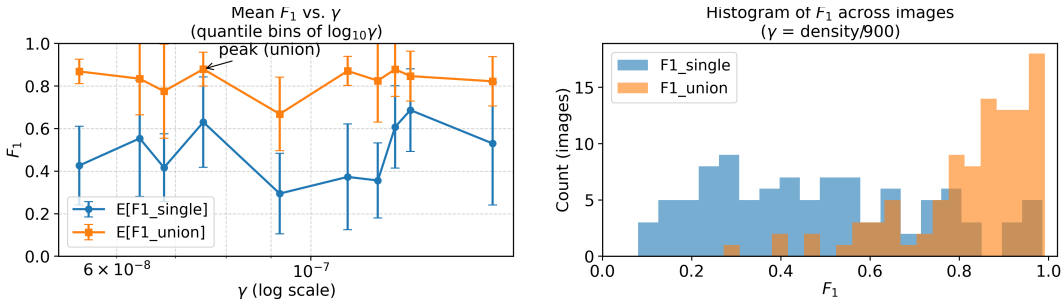
[2] Algorithm 1 presents a simplified version of the optimization process to illustrate the underlying mechanics. In practice, scanning all nodes (Line 6) results in many "ignorable checks" that do not yield strict improvements. To achieve high performance, we utilize the `community_leiden` method (Traag et al., 2019) (of the `igraph` Python library) optimizing the CPM (Traag et al., 2011), which serves as the core of our pipeline.

[3] In particular, see Algorithm 2 in Appendix E of França et al. (2026), where we also provide formal definitions, algorithmic details, and additional properties of $\text{F}_1^{\text{single}}$ and $\text{F}_1^{\text{union}}$.



(a) (b) (c) (d) (e)

Figure 3: Projections from a multi-community partition. (a) Original image. (b) Binary ground-truth mask. (c) Equilibrium partition by hedonic mechanism. (d) Best single community selected by $\mathrm{F}_1^{\text{single}}$ at $\gamma = 7.63 \times 10^{-6}$. (e) Best subset of communities selected by $\mathrm{F}_1^{\text{union}}$ at $\gamma = 2.96 \times 10^{-5}$.



(a) (b)

Figure 4: (a) Mean $\mathrm{F}_1^{\text{single}}$ and $\mathrm{F}_1^{\text{union}}$ as a function of γ (induced by $\mathrm{density}(G)/900$). (b) Global distributions of $\mathrm{F}_1^{\text{single}}$ and $\mathrm{F}_1^{\text{union}}$ over 100 images after selecting the best GT per image.

Figure 4 highlights the key insights from our experiments (see Appendix A in França et al. (2026) for additional details). Our numerical results identified $c = 900$ as the optimal value to place most instances in the fragmented-but-recoverable regime (Fig. 4(a)). Across all images, we obtain $\mathbb{E}[\mathrm{F}_1^{\text{union}}] \approx 0.828$ (median ≈ 0.868) and $\mathbb{E}[\mathrm{F}_1^{\text{single}}] \approx 0.488$, with an average gap $\mathbb{E}[\mathrm{F}_1^{\text{union}} - \mathrm{F}_1^{\text{single}}] \approx 0.340$ (cf. Fig. 4(b)). The significant average performance gap suggests that many apparent segmentation failures are actually fragmented, yet entirely recoverable, equilibria. Indeed, the partition produced by the hedonic mechanism often contains the object, but is distributed across several coalitions.

4 Pixel-Graph Construction

Recall that our pipeline is given by Figure 5. We now focus on the pixel-graph construction (Image $\rightarrow$ Graph). Each image is converted into an undirected weighted pixel graph $G = (V, E, w)$, where pixels are nodes and edges connect 8-neighborhood spatial adjacencies.

Inspired by Carreira and Sminchisescu (2011), we assign each adjacent pair (u, v) a continuous affinity that combines color similarity and boundary evidence. Let $I(u) \in [0, 255]^3$ be the RGB vector at pixel u, and let $B : \Omega \rightarrow [0, 1]$ be a normalized edge map used as a proxy for boundary probability. We define

$$w_{uv} = \exp\left(-\frac{\|I(u) - I(v)\|_2^2}{\sigma_{\text{color}}^2}\right) \exp\left(-\frac{\max\{B(u), B(v)\}}{\sigma_{\text{edge}}^2}\right). \tag{3}$$

In our implementation, B is a normalized Canny edge map (a lightweight proxy for gPb). Thus, affinities are high for similar-color neighbors and low near strong boundaries; edges with very small affinity are discarded to obtain a sparse local graph. This biases coalitions toward homogeneous regions and encourages boundaries to align with salient contours.

For the qualitative visualizations in Fig. 2(b)–(c), graphs were constructed from downsampled images to keep nodes and edges discernible. All quantitative experiments in Section 3 use the original pixel-level resolution of the Weizmann dataset, without pre-processing or scaling.

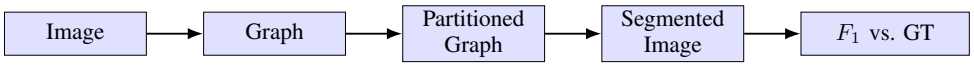


Figure 5: Diagnostic pipeline: image → graph → hedonic partition → binary projection and evaluation.

5 Conclusion

We proposed image segmentation as an interpretable testbed for studying equilibrium in hedonic games. This spatial and visual setting makes equilibrium structures directly *inspectable*, providing an intuitive way to analyze how coalition mechanisms behave. In particular, it highlights how a resolution parameter reshapes equilibrium geometry in a visually grounded domain.

The full version of this work is available at França et al. (2026), where all appendices can be found. Our results demonstrate that a density-normalized rule, $\gamma = \mathrm{density}(G)/c$, places most instances in favorable regimes by preventing extreme over- or under-fragmentation (Appendix B). Appendix C shows that a low $\mathrm{F}_1^{\text{union}} - \mathrm{F}_1^{\text{single}}$ gap reflects either mutual success for cohesive structures or intrinsic failure for detailed objects. Appendix D further confirms the robustness of the mechanism, showing no difference in equilibrium accuracy whether initialization starts from a fragmenting grand coalition or from merging isolated nodes. Finally, Appendix E formalizes the computation of these binary projections and establishes the theoretical recoverability limits of $\mathrm{F}_1^{\text{union}}$.

Future work includes evaluating alternative graph constructions and the Weizmann two-object subset. We also aim to investigate the relationship between accuracy and resolution, focusing on merging fragmented but recoverable coalitions to reduce the gap between $\mathrm{F}_1^{\text{single}}$ and $\mathrm{F}_1^{\text{union}}$.

References

Sharon Alpert, Meirav Galun, Ronen Basri, and Achi Brandt. Image segmentation by probabilistic bottom-up aggregation and CUE integration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2007.

Konstantin E Avrachenkov, Aleksei Y Kondratev, Vladimir V Mazalov, and Dmytro G Rubanov. Network partitioning algorithms as cooperative games. *Computational Social Networks*, 5(1), 2018. doi: 10.1186/s40649-018-0059-5.

Haris Aziz and Rahul Savani. Hedonic games. In *Handbook of Computational Social Choice*, pages 356–376. Cambridge University Press, Cambridge, 2016.

Anna Bogomolnaia and Matthew O Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

Joao Carreira and Cristian Sminchisescu. CPMC: Automatic object segmentation using constrained parametric min-cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7):1312–1328, 2011.

Peter Christen, David J Hand, and Nishadi Kirielle. A review of the f-measure: its history, properties, criticism, and alternatives. *ACM Computing Surveys*, 56(3):1–24, 2023.

Lucas Lopes Felipe, Konstantin Avrachenkov, and Daniel Sadoc Menasché. From Leiden to Pleasure Island: The Constant Potts Model for community detection as a hedonic game. *Physica A: Statistical Mechanics and its Applications*, page 130989, 2025.

Santo Fortunato and Marc Barthelemy. Resolution limit in community detection. *Proceedings of the national academy of sciences*, 104(1):36–41, 2007.

Pedro Henrique de Paula França, Lucas Lopes Felipe, and Daniel Sadoc Menasché. Visualizing coalition formation: From hedonic games to image segmentation. *arXiv preprint arXiv:2603.07890*, 2026.

Galun, Sharon, Basri, and Brandt. Texture segmentation by multiscale aggregation of filter responses and shape elements. In *Proceedings Ninth IEEE International Conference on Computer Vision*, pages 716–723. IEEE, 2003.

Himanshu Mittal, Avinash Chandra Pandey, Mukesh Saraswat, Sumit Kumar, Raju Pal, and Garv Modwel. A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets. *Multimedia Tools and Applications*, 81(24):35001–35026, 2022.

Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 731–737. IEEE, 1997.

Vincent A Traag, Paul Van Dooren, and Yurii Nesterov. Narrow scope for resolution-limit-free community detection. *Physical Review E*, 84(1):016114, 2011.

Vincent A Traag, Ludo Waltman, and Nees Jan van Eck. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports*, 9(1):5233, 2019.