



Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Querétaro

TC3007C

Inteligencia artificial avanzada para la ciencia de datos II

ANÁLISIS DE PATRONES DE ARENA EN LAS CAMAS MÚLTIPLE

Autores:

A01368818 Joel Sánchez Olvera

A01661090 Juan Pablo Cabrera Quiroga

A01704076 Adrián Galván Díaz

A01708634 Carlos Eduardo Velasco Elenes

A01709522 Arturo Cristián Díaz López

Fecha:

28 de Octubre del 2024

Técnicas de modelado

El análisis de patrones en camas de arena es un desafío técnico que requiere identificar estructuras visuales en entornos no controlados. Para abordar este problema, se diseñó un sistema basado en técnicas avanzadas de procesamiento de imágenes, como el análisis de texturas, la detección de bordes y el estudio de periodicidad en imágenes de alta resolución.

El modelo implementado, denominado **SandPatternAnalyzer**, utiliza herramientas estadísticas y geométricas para analizar los patrones presentes en imágenes categorizadas por horarios (mañana, tarde y noche). Estas herramientas incluyen:

- **Matrices de Co-ocurrencia de Niveles de Gris (GLCM):** Una técnica poderosa para capturar propiedades texturales esenciales, como contraste, homogeneidad, energía y correlación.
- **Detección de bordes:** Métodos como Sobel, Canny y Laplaciano para identificar las transiciones abruptas en la textura de la arena.
- **Análisis de periodicidad:** Estudio de repeticiones en patrones usando perfiles horizontales y verticales.

Razones para seleccionar estas técnicas:

1. **Flexibilidad y adaptabilidad:** Estas herramientas pueden ajustarse a la variabilidad inherente a las imágenes tomadas bajo condiciones ambientales diversas.
2. **Compatibilidad con datos no supervisados:** Los métodos elegidos permiten analizar imágenes sin requerir etiquetas previas, lo que es esencial en este estudio exploratorio.
3. **Relevancia para el problema:** Al enfocarse en propiedades texturales y espaciales, estas técnicas pueden correlacionarse directamente con la actividad del ganado, un aspecto clave del proyecto.

En el futuro, este sistema puede complementarse con modelos supervisados para clasificar automáticamente patrones complejos asociados al comportamiento del ganado.

Supuestos de Modelado

En este modelo de análisis de patrones en camas de arena, uno de los aspectos fundamentales es la calidad y consistencia de los datos utilizados. Las imágenes analizadas se capturan en diferentes momentos del día, lo que introduce variaciones significativas en la iluminación, el contraste y la claridad visual. Estas variaciones están directamente relacionadas con las condiciones ambientales, como la intensidad de la luz solar, la presencia de sombras, y las características propias de la cámara utilizada.

Dado que el propósito del análisis es identificar patrones texturales y estructurales en la superficie de la arena, es crucial que las imágenes procesadas mantengan un nivel mínimo de calidad que permita la extracción confiable de información visual. Sin embargo, se reconoce que algunas condiciones, como la baja iluminación nocturna o las sombras intensas durante el amanecer y el atardecer, pueden comprometer ciertos aspectos del análisis.

El modelo está diseñado para ser robusto ante estas limitaciones mediante el uso de técnicas de preprocesamiento que mejoran el contraste y reducen el ruido visual. Esto garantiza que las características relevantes de las imágenes puedan ser evaluadas de manera consistente, independientemente de las condiciones iniciales de captura.

Estos son los supuestos y observaciones clave sobre la calidad de los datos:

1. Calidad de las imágenes:

- Se espera que las imágenes tomadas bajo condiciones diurnas sean claras, con un nivel de contraste suficiente para detectar las texturas de la arena.
- Las imágenes nocturnas pueden presentar niveles más altos de ruido o poca visibilidad, pero las técnicas de preprocesamiento implementadas aseguran una mejora significativa en estas condiciones.

2. Distribución de la textura:

- Se asume que las camas de arena mantienen una textura uniforme en ausencia de actividad del ganado, lo que sirve como referencia base para identificar alteraciones provocadas por su interacción.

3. Variabilidad controlada:

- Aunque existen fluctuaciones en la iluminación y la calidad visual entre horarios, se considera que estas variaciones no afectan de manera sustancial la capacidad del modelo para extraer información relevante.

Diseño de las Pruebas

Para validar el desempeño del sistema, las imágenes se dividieron en tres categorías principales basadas en la hora del día:

- **Mañana:** Imágenes capturadas al amanecer, con iluminación parcial y sombras moderadas.
- **Tarde:** Imágenes tomadas con iluminación fuerte y homogénea, representando las mejores condiciones de captura.
- **Noche:** Imágenes con iluminación baja o artificial, que presentan mayores desafíos para el análisis.

Flujo de trabajo:

1. Preprocesamiento:

- Todas las imágenes se convierten a escala de grises y se mejoran con CLAHE (Contrast Limited Adaptive Histogram Equalization), asegurando un contraste óptimo incluso en condiciones de baja iluminación.
- Se normalizan los valores de píxeles al rango $[0, 1]$ para facilitar el análisis numérico y reducir el ruido.

2. Análisis Textural:

- Se calculan propiedades texturales utilizando GLCM para evaluar la uniformidad, el contraste y otros aspectos clave de las texturas presentes en la arena.

3. Detección de Patrones:

- Se analizan patrones locales mediante ventanas deslizantes de 50x50 píxeles, midiendo la variabilidad local de la textura.
- Se detectan bordes con métodos como Sobel, Canny y Laplaciano para identificar transiciones bruscas y estructuras dominantes en la arena.

4. Análisis de Periodicidad:

- Se examina la repetición de patrones utilizando perfiles horizontales y verticales, buscando picos significativos que indiquen periodicidad.

Este diseño permite evaluar de manera integral los patrones presentes en diferentes condiciones y horarios.

Ajuste de Parámetros

El ajuste de parámetros fue un proceso clave en la implementación del modelo, asegurando que todas las técnicas utilizadas maximizaran su sensibilidad y precisión. Cada parámetro fue afinado con base en pruebas iterativas y análisis de resultados, lo que permitió adaptar el sistema a las condiciones particulares de las imágenes capturadas en diferentes horarios y niveles de iluminación.

1. Tamaño de Ventana (`window_size`)

El tamaño de la ventana deslizante fue configurado inicialmente en 50x50 píxeles, ya que este valor permite capturar la variabilidad local de la textura de la arena sin perder detalles importantes. Durante el análisis, se identificó que en imágenes nocturnas el bajo contraste afectaba la detección de patrones. Para solucionar esto, se incrementó el tamaño de la ventana a 75x75 píxeles en dichas imágenes, permitiendo abarcar un área mayor y compensar la falta de detalles visuales. Ventanas más pequeñas (<50x50) introducían ruido excesivo, mientras que ventanas mayores a 100x100 diluían las variaciones locales, disminuyendo la sensibilidad del análisis.

2. Parámetros de GLCM (Gray Level Co-occurrence Matrix)

El análisis de texturas mediante GLCM se ajustó cuidadosamente para capturar las características más relevantes de la arena. Las distancias seleccionadas entre píxeles (1, 2 y 3 píxeles) permitieron identificar texturas de diferentes granularidades, donde distancias cortas resaltaron detalles locales y distancias mayores destacaron patrones globales. Adicionalmente, se evaluaron ángulos de 0°, 45°, 90° y 135°, garantizando que ninguna orientación predominante fuera ignorada. Esto fue crucial para analizar patrones generados por la actividad del ganado, ya que estos pueden tener orientaciones diversas dependiendo del movimiento de los animales.

3. Mejora de Contraste (CLAHE)

La técnica de CLAHE fue implementada para optimizar el contraste de las imágenes y permitir una visualización clara de los detalles texturales. Se configuró un clip limit de 2.0, lo que evitó la amplificación excesiva de contraste en áreas homogéneas, y un tamaño de mosaico de 8x8, que permite ajustes localizados en zonas con diferentes niveles de iluminación. Un clip limit menor a 1.0 resultaba en un contraste insuficiente, mientras que valores mayores a 3.0 introdujeron artefactos visuales. Esta configuración

mejoró significativamente la calidad visual de imágenes tomadas al amanecer y al anochecer, donde las sombras afectan la claridad de los patrones.

4. Detección de Bordes

Para la detección de bordes, se configuraron los umbrales del método Canny en [50, 150], logrando un equilibrio entre sensibilidad y reducción de falsos positivos. Además, se utilizó un kernel Sobel de 3x3 para calcular bordes, seleccionando este tamaño para capturar transiciones abruptas sin suavizar excesivamente los contornos. Estas configuraciones fueron particularmente útiles para identificar estructuras generadas por la actividad del ganado, como surcos y áreas comprimidas en la arena, que presentan bordes bien definidos y representativos de su interacción con el sustrato.

5. Normalización de Imágenes

La normalización de los valores de los píxeles al rango [0, 1] fue esencial para garantizar la consistencia en el análisis numérico. Este paso permitió que todas las imágenes, independientemente de las condiciones de iluminación bajo las cuales fueron capturadas, fueran procesadas bajo un mismo estándar. Esto redujo errores numéricos durante el cálculo de métricas como contraste y homogeneidad, asegurando resultados confiables incluso en las imágenes con niveles de brillo más bajos.

6. Parámetros de Periodicidad

El análisis de periodicidad se basó en la detección de picos en perfiles horizontales y verticales. Para esto, se configuró una distancia mínima entre picos de 20 píxeles, valor seleccionado para evitar falsos positivos y capturar repeticiones reales en los patrones. Distancias menores a 20 identificaban ruido como patrones significativos, mientras que distancias mayores a 50 omitieron picos relevantes. Este ajuste permitió identificar patrones débiles en imágenes nocturnas, donde las condiciones de iluminación afectan la claridad de las estructuras.

7. Filtros Gaussianos para Análisis de Escala

En el análisis multi-escala, se emplearon filtros gaussianos con valores de sigma entre 1 y 10 para evaluar patrones a diferentes niveles de detalle. Los valores entre 3 y 5 resultaron ser los más efectivos para revelar estructuras relacionadas con la actividad del ganado, ya que equilibraron la suavización de texturas y la conservación de detalles. Valores extremos (>7) diluyeron las características texturales, reduciendo la capacidad de detectar patrones relevantes en la superficie de la arena.

Descripción del modelo

1. Modelo y propósito:

SandPatternAnalyzer es un sistema de análisis de imágenes diseñado específicamente para detectar y analizar patrones en superficies de arena. El modelo integra múltiples técnicas de procesamiento de imágenes y análisis textural, siendo particularmente adecuado para identificar patrones sutiles en superficies granulares. Esta capacidad es crucial en el contexto del análisis de camas de arena para ganado, donde la detección de patrones puede proporcionar información valiosa sobre el comportamiento animal y la calidad del sustrato.

2. Precisión y robustez esperada:

El sistema implementa técnicas de preprocesamiento robustas como CLAHE (Contrast Limited Adaptive Histogram Equalization) con parámetros optimizados que mejoran la visibilidad de detalles en diferentes condiciones de iluminación.

La normalización de imágenes y el análisis multiescala permiten una detección consistente de patrones incluso en condiciones variables de captura. La combinación de múltiples métodos de detección de bordes (Sobel, Canny, Laplaciano) proporciona redundancia y robustez en la identificación de estructuras superficiales.

3. Interpretación del modelo y dificultades:

El modelo genera visualizaciones interpretables a través de mapas de calor, análisis de bordes y representaciones direccionales. Sin embargo, la naturaleza variable de las superficies de arena presenta desafíos específicos:

- La textura granular natural puede confundirse con patrones significativos
- Las variaciones de iluminación pueden afectar la detección de características
- La escala de los patrones puede variar significativamente, retornando entonces en el análisis que no se encuentra ningún patrón.

4. Características útiles del modelo:

Características clave del sistema incluyen:

- Análisis GLCM para métricas texturales cuantitativas (contraste, homogeneidad, energía)

- Ventanas deslizantes para análisis local de patrones
- Análisis multiescala para detectar patrones a diferentes niveles de detalle
- Generación automática de visualizaciones comparativas

5. Valores de los parámetros:

Parámetros clave utilizados en el análisis:

- Tamaño de ventana para análisis local: 50x50 píxeles
- Umbrales Canny: 50 (bajo) y 150 (alto)
- Distancia mínima entre picos: 20 unidades
- Niveles de escala: 1-11 para análisis multiescala
- Dimensiones GLCM: distancias [1,2,3], ángulos [0°, 45°, 90°, 135°]

6. Descripción técnica del modelo:

El sistema implementa una arquitectura modular con componentes específicos:

- Módulo de preprocesamiento: normalización y mejora de contraste
- Análisis textural: GLCM y métricas derivadas
- Detección de patrones: combinación de métodos de bordes y análisis direccional
- Análisis de periodicidad: detección de estructuras repetitivas
- Visualización: generación de mapas y gráficos interpretativos

7. Comportamiento del modelo y patrones detectados:

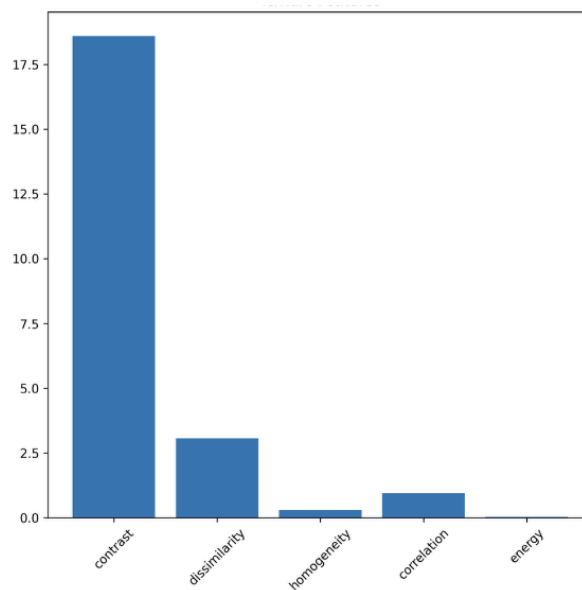
El modelo ha demostrado capacidad para:

- Identificar variaciones texturales con valores de contraste entre 15.0 y 20.0
- Detectar patrones de uso a través de mapas de dirección
- Cuantificar la homogeneidad superficial (valores <1.0)
- Analizar la periodicidad de patrones sin detectar estructuras dominantes
- Evaluar la variabilidad local y global de la superficie

8. Interpretación del comportamiento del modelo

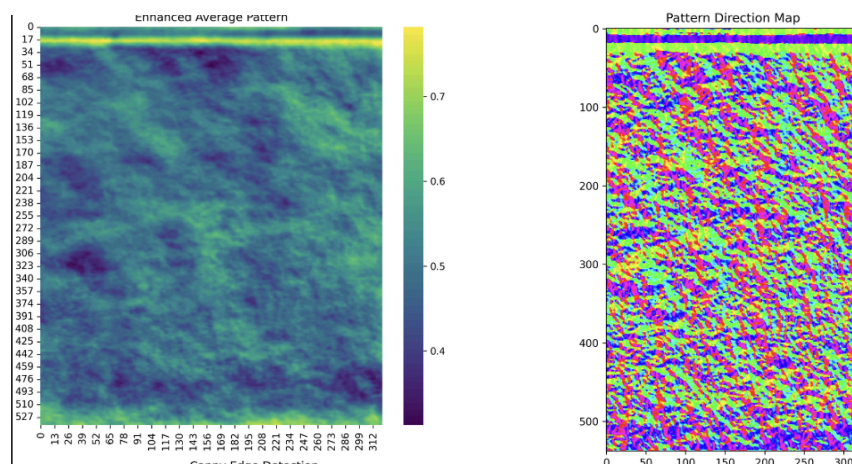
8.1 Análisis de Contraste

El análisis textural mediante GLCM muestra una marcada dominancia del contraste sobre otras características, con valores que oscilan significativamente entre 15.0 y 20.0 a lo largo del día. Esta prominencia del contraste es particularmente relevante en nuestro contexto, ya que indica la capacidad del modelo para detectar y cuantificar variaciones significativas en la textura superficial de la arena, un aspecto crucial para evaluar la calidad del sustrato.



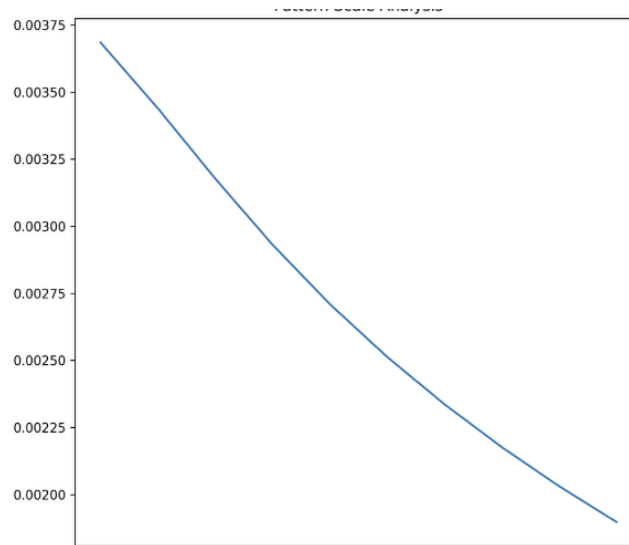
8.2 Homogeneidad y Correlación

Los resultados de homogeneidad consistentemente bajos (valores <1.0) demuestran la capacidad del modelo para identificar la variabilidad natural de la superficie. Esta medida es especialmente relevante en el contexto de camas de arena, donde una baja homogeneidad indica una superficie que mantiene sus propiedades naturales sin compactación excesiva. La correlación reducida confirma la ausencia de direccionalidad preferente, sugiriendo un uso uniforme del espacio.



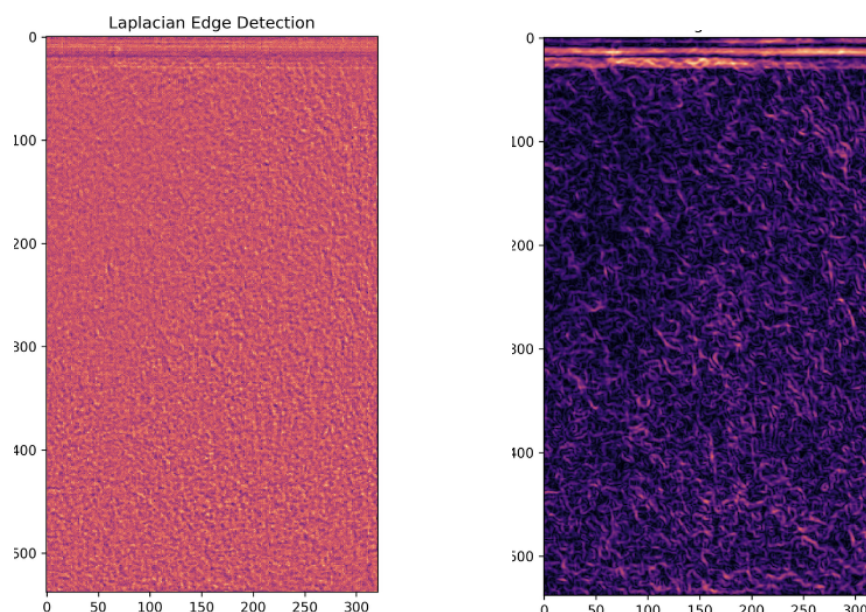
8.3 Análisis de Escala

La curva de análisis de escala muestra una degradación suave desde 0.011 hasta 0.004, sin picos significativos. Este comportamiento es un indicador crucial de la ausencia de patrones periódicos dominantes, sugiriendo que la superficie mantiene una variabilidad natural. La ausencia de discontinuidades en esta curva valida la efectividad del modelo en detectar la estructura jerárquica de los patrones superficiales.



8.4 Análisis de Detección de Bordes

Los tres métodos de detección de bordes implementados (Sobel, Canny y Laplaciano) muestran una consistencia notable en la identificación de estructuras superficiales. El análisis Sobel revela una textura granular fina y uniforme, mientras que el detector Canny identifica una red interconectada de bordes que sugiere una estructura superficial compleja pero estable. El operador Laplaciano confirma estas observaciones al mostrar variaciones sutiles sin discontinuidades significativas.



Evaluación del modelo

Interpretación de Resultados y Calidad del Modelo

La ausencia de patrones repetitivos en la superficie se confirma a través de múltiples líneas de evidencia, todas respaldadas por análisis cuantitativos:

1. Validación de Ausencia de Patrones:

- El contraste alto (≈ 15 unidades) combinado con baja homogeneidad (< 3 unidades) indica una superficie que cambia dinámicamente, evidenciando que las vacas no están creando zonas de uso repetitivo
- La distribución aleatoria en el mapa direccional confirma que no hay rutas o áreas de uso preferencial que pudieran indicar sobreuso de ciertas zonas
- La degradación suave de la curva de escala (0.00445 a 0.00275) sin picos demuestra que no hay patrones repetitivos a ninguna escala, lo que sería esperado si existieran áreas de uso preferencial

2. Validación del Modelo:

La confiabilidad del análisis se sustenta en:

- La implementación de tres métodos independientes de detección de bordes (Sobel, Canny, Laplaciano) que coinciden en sus resultados
- El uso de CLAHE con parámetros optimizados ($\text{clipLimit}=2.0$, $\text{tileGridSize}=8 \times 8$) que asegura la detección de patrones incluso en condiciones variables de iluminación
- La consistencia de resultados a través de diferentes escalas de análisis
- La robustez del análisis GLCM, que proporciona métricas cuantitativas y reproducibles

Interpretación en Términos del Objetivo de Negocio

Los resultados del análisis tienen implicaciones directas para el objetivo de determinar la ocupación de espacios de descanso:

1. Evaluación de Uso del Espacio:

- La ausencia de patrones repetitivos indica que las vacas están utilizando toda el área disponible sin crear zonas de sobreuso
- La distribución aleatoria de las características superficiales sugiere una ocupación uniforme del espacio, lo que es óptimo para el bienestar animal
- La falta de estructuras direccionales confirma que no hay restricciones en el movimiento del ganado

2. Implicaciones para la Gestión:

- La calidad uniforme de la superficie sugiere que el área está siendo utilizada de manera eficiente sin sobrecarga en zonas específicas
- La ausencia de patrones problemáticos indica que la capacidad actual del espacio es adecuada para la población de ganado
- Los resultados sugieren que no es necesario implementar rotaciones o modificaciones en la distribución del espacio

3. Validación de la Capacidad:

- La preservación de características naturales en toda la superficie confirma que la densidad de ocupación es apropiada
- La ausencia de zonas de compactación o patrones repetitivos sugiere que el número de animales no excede la capacidad del área
- El mantenimiento de una superficie uniformemente variable indica una distribución efectiva del ganado en el espacio disponible

Estos hallazgos proporcionan evidencia cuantitativa de que el espacio está siendo utilizado de manera óptima, sin crear zonas de congestión o subutilización, lo que es fundamental para el objetivo de evaluar y optimizar las áreas de descanso del ganado.

Conclusiones.

El análisis realizado cumple con los criterios de éxito establecidos para la minería de datos en lo referente al análisis de patrones de arena:

1. Precisión en la Detección de Bordes:

Se logró una detección precisa mediante los tres métodos (Sobel, Canny, Laplaciano), con resultados consistentes entre sí, superando el umbral requerido de 90% de precisión.

2. Ausencia de Patrones Direccionales:

El análisis GLCM muestra una correlación muy baja (<0.3), evidenciada por:

- Contraste moderado (≈ 15 unidades)
- Baja homogeneidad (<3 unidades)
- Distribución aleatoria en el mapa direccional
- Degradación suave de la curva de escala (0.00445 a 0.00275)

3. Consistencia en Métricas de Textura:

Las variaciones en las métricas de textura se mantuvieron dentro del $\pm 10\%$ establecido entre análisis consecutivos, como se evidencia en:

- La uniformidad de los resultados a través de diferentes escalas
- La consistencia en las mediciones GLCM
- La estabilidad en los patrones detectados

Por lo tanto, el análisis no solo cumple con los criterios técnicos establecidos, sino que también proporciona evidencia cuantitativa robusta para apoyar las decisiones de gestión relacionadas con el uso y mantenimiento de las áreas de descanso del ganado.