

**Nombre del proyecto**

|  |
| --- |
| **INVESTIGACIÓN Y COMPARACIÓN EXHAUSTIVA DE TÉCNICAS DE IA PARA**  **CLASIFICACIÓN DE IMÁGENES TIROIDEAS** |

|  |  |  |  |
| --- | --- | --- | --- |
| **Profesor** | **Materia** | **Fecha de Inicio** | **Fecha de Finalización** |
| Ing. Gladys Villegas | Proyecto Integrador | 22/09/2025 | 24/09/2025 |

**Resumen del Proyecto**

|  |
| --- |
| En este trabajo se realizó una investigación exhaustiva de técnicas de inteligencia artificial aplicadas a la clasificación de imágenes tiroideas. Se analizaron cinco enfoques principales: Random Forest, Support Vector Machines (SVM), Redes Neuronales Convolucionales (CNN), Vision Transformers (ViT) y Transfer Learning con CNN.  Para cada técnica se revisaron sus fundamentos teóricos, ventajas, limitaciones, complejidad computacional, casos de uso médicos e industriales y finalmente su aplicabilidad al proyecto.  En conclusión, el proyecto se enfocará en CNN con transfer learning como solución óptima, complementado por CNN tradicionales y Vision Transformers para comparación y validación experimental. |

**Responsables del proyecto**

|  |  |  |
| --- | --- | --- |
| **Nombre** | **Rol** | **Tarea** |
| Christian García | Estudiante/Desarrollador | Análisis Comparativo de Algoritmos |
| Byron Piedra | Estudiante/ Scrum Master | Análisis Comparativo de Algoritmos |

**INVESTIGACIÓN Y COMPARACIÓN EXHAUSTIVA DE TÉCNICAS DE IA PARA**

**CLASIFICACIÓN DE IMÁGENES TIROIDEAS**

**1. RANDOM FOREST**

**1.1 Descripción Teórica Profundizada**

**Fundamentos matemáticos avanzados:**

* Ensemble method basado en el Teorema del Límite Central aplicado a árboles de decisión
* Bootstrap aggregating (bagging) + random feature selection
* Impurity measures: Gini index ≠ H(X) = -Σp(x)log p(x) para división de nodos

**Mecanismo de funcionamiento detallado:**

1. Creación de múltiples subsets via bootstrapping (sampling con reemplazo)
2. Entrenamiento de árboles independientes con √n features aleatorias
3. Aggregation por votación mayoritaria (clasificación) o promedio (regresión)
4. Out-of-bag error estimation para validación interna

**Parámetros críticos y su impacto:**

* n\_estimators: 100-500 (ley de rendimientos decrecientes >200)
* max\_depth: control fundamental de overfitting
* min\_samples\_split: 2-5 para datos médicos
* max\_features: 'sqrt' ó 'log2' para alta dimensionalidad

**1.2 Análisis Profundo de Ventajas/Desventajas**

**Fortalezas técnicas avanzadas:**

* **Robustez estadística:** Reducción de varianza mediante promediado
* **Handling missing values:** Mecanismos internos de imputación
* **Feature importance:** Permite interpretabilidad clínica
* **Resistencia a outliers:** Medianas más robustas que promedios

**Limitaciones técnicas críticas:**

* **Curse of dimensionality:** Performance decae con >100 features
* **Correlation bias:** Tiende a features correlacionados
* **Extrapolation poor:** No generaliza bien fuera del rango de entrenamiento
* **Memory footprint:** O(n\_estimators × 2^max\_depth)

**1.3 Complejidad Computacional Detallada**

**Análisis Big-O riguroso:**

* **Entrenamiento:** O(n\_estimators × m × n log n)
  + m = número de características, n = muestras
  + Árbol individual: O(m × n log n)
* **Predicción:** O(n\_estimators × depth)
* **Memoria:** O(n\_estimators × 2^depth × m)

**Optimizaciones prácticas:**

* Parallelization: O(n\_estimators/k) con k cores
* Incremental learning posible pero limitado
* GPU acceleration mínima beneficio

**1.4 Casos de Uso Expandidos con Ejemplos Reales**

**Implementaciones médicas documentadas:**

1. **Mayo Clinic:** Predicción de riesgo cardiovascular con 150+ variables clínicas
2. **MD Anderson Cancer Center:** Clasificación de subtipos de cáncer basado en biomarkers
3. **Stanford Medicine:** Detección de sepsis en UCIs usando datos vitales temporales

**Industrias y aplicaciones actuales:**

* **Fintech:** Fraud detection con datos transaccionales
* **Retail:** Customer segmentation y recommendation engines
* **Manufacturing:** Predictive maintenance de equipos
* **Agricultura:** Yield prediction basado en condiciones ambientales

**1.5 Aplicabilidad al Proyecto Tiroides**

**Análisis de idoneidad técnica:**

* **Ventaja:** Interpretabilidad para validación clínica
* **Limitación crítica:** Requiere feature engineering manual de imágenes
* **Riesgo:** Pérdida de información espacial en ultrasonidos

**Requisitos de implementación:**

* Extracción de características: Texture, shape, echogenicity features
* Dimensionality reduction: PCA de 224×224×3 → ~100 features
* Validación clínica de features seleccionadas

**Decisión final:** **EXCLUIR** - Inviable para clasificación directa de imágenes raw

**2. SUPPORT VECTOR MACHINES (SVM)**

**2.1 Descripción Teórica Profundizada**

**Fundamentos matemáticos avanzados:**

* Optimización convexa: min(1/2||w||²) sujeto a y\_i(w·x\_i + b) ≥ 1
* Dual problem: max(Σα\_i - 1/2ΣΣα\_iα\_jy\_iy\_jK(x\_i,x\_j))
* Kernel trick: φ(x) mapping implícito a espacio de alta dimensión

**Kernels especializados para imágenes:**

* RBF: K(x,y) = exp(-γ||x-y||²) para similitud local
* Polynomial: Captura interacciones entre píxeles
* Histogram intersection: Especializado para características de imagen

**Parámetros de fine-tuning:**

* C: Trade-off margin-error (1-10 para datos médicos)
* γ en RBF: Controla influencia de cada muestra (auto-scale recomendado)
* Kernel selection: Validación cruzada nested

**2.2 Análisis Profundo de Ventajas/Desventajas**

**Fortalezas teóricas únicas:**

* **Margen máximo:** Generalización garantizada teóricamente
* **Kernel flexibility:** Adaptable a diversos tipos de datos
* **Sparsity:** Solo support vectors afectan decisiones
* **Convexidad:** Solución global garantizada

**Limitaciones prácticas severas:**

* **O(n³) complexity:** Inviable para n > 10,000
* **Kernel selection heuristic:** No systematic approach
* **Memory bottleneck:** O(n\_sv × d) storage
* **Multi-class complexity:** One-vs-one/one-vs-all overhead

**2.3 Complejidad Computacional Detallada**

**Análisis de escalabilidad real:**

* **Entrenamiento:** O(n³) worst-case, O(n²) promedio con optimizations
* **SMO algorithm:** O(n² × d) practical implementation
* **Predicción:** O(n\_sv × d) por instancia
* **Memory:** O(n\_sv × d) para model storage

**Límites prácticos de dataset size:**

* Máximo viable: ~50,000 muestras en cluster HPC
* Óptimo: 1,000-10,000 muestras para single machine
* GPU acceleration limitada por naturaleza algorítmica

**2.4 Casos de Uso Expandidos con Ejemplos Reales**

**Implementaciones médicas de alto impacto:**

1. **MIT Clinical ML Group:** Clasificación de tejidos en histopatología con kernel personalizado
2. **Johns Hopkins Radiology:** Detección de microcalcificaciones en mamografías
3. **NIH Bioinformatic Center:** Análisis de expresión génica para diagnóstico cáncer

**Aplicaciones industriales especializadas:**

* **Bioinformática:** Protein structure prediction
* **Finanzas de alta frecuencia:** Time-series classification
* **Geospatial imaging:** Land use classification from satellite data
* **Quality control:** Defect detection in manufacturing

**2.5 Aplicabilidad al Proyecto Tiroides**

**Viabilidad técnica assessment:**

* **Potencial:** High-dimensional separation con kernel RBF
* **Limitación:** n=1,372 imágenes → O(1,372³) ≈ 2.5×10⁹ operaciones
* **Riesgo:** Kernel selection crítica para texturas médicas

**Requisitos computacionales:**

* Cluster HPC para entrenamiento práctico
* Feature extraction necesaria para dimensionalidad
* Cross-validation extensiva para hyperparameter tuning

**Decisión final:** **EXCLUIR** - Complexity computacional prohibitiva

**3. REDES NEURONALES CONVOLUCIONALES (CNN)**

**3.1 Descripción Teórica Profundizada**

**Fundamentos matemáticos avanzados:**

* Convolución discreta: (f∗g)[n] = Σf[m]g[n-m] con padding y stride
* Non-linearities: ReLU, Leaky ReLU, ELU activations
* Backpropagation: ∇W(l) = δ(l+1) ⋅ a(l)^T con chain rule
* Batch normalization: γ(x-μ)/√(σ²+ε) + β para estabilidad

**Arquitecturas especializadas médicas:**

* U-Net variants para segmentación
* DenseNet connections para feature reuse
* Attention mechanisms para regions de interés
* 3D CNNs para volumetric data

**Hyperparameter optimization space:**

* Learning rate: 1e-3 to 1e-5 con schedulers
* Batch size: 16-128 dependiendo de GPU memory
* Optimizers: AdamW > Adam > SGD con momentum
* Regularization: Dropout 0.2-0.5, L2 weight decay

**3.2 Análisis Profundo de Ventajas/Desventajas**

**Fortalezas arquitectónicas únicas:**

* **Translation invariance:** Patrones reconocibles en cualquier posición
* **Hierarchical feature learning:** Low-level → high-level features
* **Parameter sharing:** Efficiency computacional radical
* **End-to-end learning:** Minimal manual feature engineering

**Desafíos prácticos significativos:**

* **Data hunger:** 1,000+ samples per class ideal
* **Overfitting risk:** Especialmente con datos limitados
* **Black box nature:** Interpretabilidad clínica desafiant
* **Hardware requirements:** GPUs esenciales para entrenamiento práctico

**3.3 Complejidad Computacional Detallada**

**Análisis layer-by-layer:**

* **Convolution layer:** O(h × w × c\_in × c\_out × k² × batch\_size)
* **Fully-connected:** O(input\_size × output\_size)
* **Total training:** O(epochs × batches × Σlayer\_complexity)

**Memory footprint analysis:**

* Activations: O(batch\_size × Σlayer\_activations)
* Parameters: O(Σlayer\_parameters)
* Gradients: Mismo order que parameters

**Scalability en práctica:**

* Single GPU: Hasta ~10M parámetros eficientemente
* Multi-GPU: Linear scaling con data parallelism
* TPU: Optimizado para convoluciones específicamente

**3.4 Casos de Uso Expandidos con Ejemplos Reales**

**Implementaciones médicas revolucionarias:**

1. **Google Health - Diabetic Retinopathy:** Screening a escala con 94%+ accuracy
2. **Mass General Hospital - Brain Tumor Classification:** Integrado en flujo clínico radiológico
3. **Stanford ML Group - Pneumonia Detection:** FDA-approved para uso clínico
4. **DeepMind - Breast Cancer Screening:** Reducción de falsos positivos en 5.7%

**Adopción industrial masiva:**

* **Automotriz:** Autonomous driving vision systems
* **Seguridad:** Facial recognition y video analytics
* **Retail:** Visual search y inventory management
* **Agricultura:** Crop monitoring y yield optimization
* **Manufactura:** Defect detection automatizado

**3.5 Aplicabilidad al Proyecto Tiroides**

**Análisis de adecuación técnica completa:**

* **Match perfecto:** Arquitectura diseñada específicamente para imágenes
* **Proven track record:** Estado del arte en imágenes médicas
* **Transfer learning:** Modelos pre-entrenados disponibles

**Requisitos específicos del proyecto:**

* Data augmentation agresiva para dataset limitado
* Fine-tuning de arquitecturas pre-entrenadas (ImageNet)
* Regularización robusta (dropout, early stopping)

**Decisión final:** **INCLUIR COMO BASELINE** - Estándar oro para comparación

**4. VISION TRANSFORMERS (ViT)**

**4.1 Descripción Teórica Profundizada**

**Fundamentos matemáticos avanzados:**

* Self-attention: Attention(Q,K,V) = softmax(QK^T/√d\_k)V
* Multi-head mechanism: Concatenación de h attention heads
* Patch embedding: División de imagen en patches de 16×16
* Positional encoding: Información espacial via embeddings aprendidos

**Arquitecturas médicas especializadas:**

* TransUNet: Transformer + U-Net para segmentación
* Swin Transformers: Hierarchical feature maps
* Medical Transformer: Attention gates para regions de interés
* 3D ViT: Para volumetric medical imaging

**Training strategies avanzadas:**

* Self-supervised pre-training: MAE, DINO, iBOT
* Knowledge distillation: Teacher-student frameworks
* Multi-scale processing: Pyramid architectures

**4.2 Análisis Profundo de Ventajas/Desventajas**

**Ventajas revolucionarias:**

* **Global context:** Atención a toda la imagen simultáneamente
* **Scalability:** Mejora con datos y modelo size
* **Transfer learning:** Pre-training en datasets masivos
* **Multi-modal fusion:** Integración con texto clínico

**Desafíos técnicos significativos:**

* **Data efficiency:** Requiere >100K imágenes para pre-training efectivo
* **Computational cost:** O(n²) en número de patches
* **Memory intensive:** Attention matrices grandes
* **Training instability:** Sensible a hyperparameters

**4.3 Complejidad Computacional Detallada**

**Análisis de bottlenecks:**

* **Self-attention:** O(n² × d) donde n = número de patches
* **MLP layers:** O(n × d²) por layer
* **Total por layer:** O(n² × d + n × d²)

**Optimizaciones prácticas:**

* Linear attention approximations: O(n × d)
* Sparse attention patterns
* Mixed-precision training
* Gradient checkpointing

**Hardware requirements realistas:**

* ViT-Base: ~86M parámetros, 16GB GPU mínimo
* ViT-Large: ~307M parámetros, múltiple GPUs necesarias
* Inference optimizations via quantization

**4.4 Casos de Uso Expandidos con Ejemplos Reales**

**Implementaciones médicas de vanguardia:**

1. **Microsoft Research - COVID-19 Diagnosis:** ViT con 98% accuracy en detección
2. **NVIDIA CLARA - Pathology Analysis:** Análisis de slides completos
3. **IBM Watson Health - Medical Imaging Analytics:** Plataforma enterprise-scale
4. **MIT J-Clinic - Cancer Diagnostics:** Integración multi-modal

**Industrias de adopción temprana:**

* **Tecnología:** Google, Meta, Microsoft para visión artificial
* **Salud digital:** Startups de diagnóstico asistido
* **Investigación farmacéutica:** Drug discovery y biomarker identification
* **Telemedicina:** Plataformas de análisis remoto

**4.5 Aplicabilidad al Proyecto Tiroides**

**Assessment de viabilidad realista:**

* **Potencial alto:** Captura contexto global en ultrasonidos
* **Limitación crítica:** Dataset de 1,372 imágenes insuficiente para training from scratch
* **Solución:** Transfer learning de modelos pre-entrenados

**Estrategia de implementación práctica:**

* Fine-tuning de ViT-Base pre-entrenado en ImageNet-21K
* Data augmentation agresiva (mixup, cutmix)
* Gradient accumulation para batch sizes efectivos

**Decisión final:** **INCLUIR COMO TÉCNICA AVANZADA** - Alto potencial con transfer learning

**5. TRANSFER LEARNING CON CNN (TÉCNICA PRINCIPAL)**

**5.1 Descripción Teórica Profundizada**

**Fundamentos teóricos avanzados:**

* Domain adaptation theory: Minimizar discrepancy entre source y target domains
* Feature reuse: Activaciones de capas iniciales son generalizables
* Fine-tuning strategies: Differential learning rates por capas

**Metodologías especializadas médicas:**

* Medical transfer learning: Pre-training en datasets médicos grandes
* Progressive resizing: Incremental image size durante training
* Discriminative learning rates: Capas iniciales → lr bajo, finales → lr alto

**Frameworks prácticos:**

* TensorFlow Hub: Modelos pre-entrenados medical imaging
* MONAI: Medical Open Network for AI
* NVIDIA NIM: Optimized medical AI models

**5.2 Análisis Profundo de Ventajas/Desventajas**

**Ventajas decisivas para el proyecto:**

* **Data efficiency:** Excelente performance con <1,000 muestras
* **Time efficiency:** Training en horas vs semanas
* **Proven effectiveness:** 95%+ problemas prácticos de visión
* **Robustness:** Menor overfitting que training from scratch

**Consideraciones prácticas críticas:**

* **Domain gap:** ImageNet → medical images discrepancy
* **Architecture constraints:** Limitado a arquitecturas pre-entrenadas
* **Hyperparameter sensitivity:** Learning rates críticos

**5.3 Complejidad Computacional Detallada**

**Fine-tuning complexity analysis:**

* **Feature extraction:** O(batches × forward\_pass) - solo inferencia
* **Fine-tuning:** O(epochs × batches × (forward + backward))
* **Memory:** Activations + gradients para capas descongeladas

**Comparative efficiency:**

* vs Training from scratch: 10-100× más rápido
* vs Feature extraction: Mejor accuracy, costo moderado
* Optimal trade-off: Fine-tuning de últimas capas

**5.4 Casos de Uso Expandidos con Ejemplos Reales**

**Implementaciones médicas masivamente exitosas:**

1. **CheXNet (Stanford):** Detección de neumonía en rayos X - 0.81 AUC
2. **Skin Cancer Detection (Google):** Clasificación de lesiones dérmicas
3. **Retina Diagnosis (DeepMind):** Detección de enfermedades oculares
4. **COVID-Net (DarwinAI):** Diagnóstico de COVID en radiografías torácicas

**Adopción industrial universal:**

* **Salud digital:** Todas las principales plataformas de telemedicina
* **Dispositivos médicos:** Integration en ultrasound machines modernas
* **Diagnóstico asistido:** Software radiológico comercial (GE, Siemens)
* **Investigación clínica:** Estudios validados multicéntricos

**5.5 Aplicabilidad al Proyecto Tiroides**

**Análisis de optimalidad técnica:**

* **Match perfecto:** Dataset limitado (1,372 imágenes)
* **Timeline compatible:** 4 semanas realista
* **Recursos adecuados:** GPU single suficiente
* **Validation clínica:** Interpretabilidad via CAM/Grad-CAM

**Plan de implementación detallado:**

1. **Semana 1:** Feature extraction con modelos pre-entrenados
2. **Semana 2:** Fine-tuning progresivo de capas
3. **Semana 3:** Optimización hyperparameters
4. **Semana 4:** Validación y interpretabilidad

**Modelos específicos recomendados:**

* EfficientNet-B3: Balance accuracy/efficiency
* DenseNet-201: Feature reuse máximo
* ResNet-50: Estabilidad comprobada

**Decisión final:** **TÉCNICA PRINCIPAL** - Óptima para todas las constraints del proyecto

**COMPARATIVA FINAL Y RECOMENDACIÓN ESTRATÉGICA**

1. **Matriz de Decisión Multicriterio:**

| **Criterio** | **Random Forest** | **SVM** | **CNN From Scratch** | **Vision Transformer** | **Transfer Learning CNN** |
| --- | --- | --- | --- | --- | --- |
| **Performance** | 2/5 | 3/5 | 4/5 | 5/5 | **5/5** |
| **Data Efficiency** | 3/5 | 2/5 | 2/5 | 1/5 | **5/5** |
| **Training Time** | 4/5 | 2/5 | 2/5 | 1/5 | **5/5** |
| **Hardware Requirements** | 5/5 | 3/5 | 3/5 | 2/5 | **4/5** |
| **Interpretability** | 5/5 | 4/5 | 2/5 | 2/5 | **3/5** |
| **4-Week Viability** | 2/5 | 2/5 | 3/5 | 2/5 | **5/5** |
| **Clinical Validation** | 5/5 | 4/5 | 3/5 | 3/5 | **4/5** |

1. **Recomendación Final Estratégica:**
   1. **Técnica Principal: Transfer Learning con CNN**

* Rationale: Maximiza performance dentro de constraints prácticos
* Architectura: EfficientNet-B3 + discriminative fine-tuning
* Timeline: Perfectamente viable en 4 semanas
  1. **Técnicas Comparativas:**
* CNN From Scratch: Baseline para validación de aproximación
* Vision Transformer: Análisis de técnicas state-of-the-art
  1. **Exclusión Justificada:**
* Random Forest/SVM: No adecuados para imagen raw
  1. **Plan de Implementación Faseada:**

**Fase 1 (Semanas 1-2):** Transfer Learning CNN - Desarrollo principal  
**Fase 2 (Semana 3):** CNN From Scratch - Validación comparativa  
**Fase 3 (Semana 4):** Vision Transformer - Análisis avanzado  
**Fase 4 (Semana 4):** Evaluación integral y conclusions